

PCT

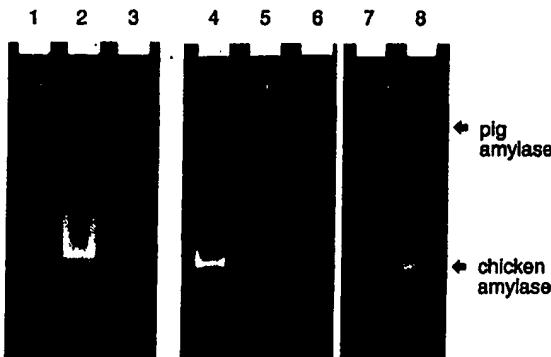
WORLD INTELLECTUAL PROPERTY ORGANIZATION  
International Bureau



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification <sup>6</sup> :	A1	(11) International Publication Number:	WO 99/43795
C12N 15/00, A01K 67/027, C12N 15/86, 9/72		(43) International Publication Date:	2 September 1999 (02.09.99)
(21) International Application Number:	PCT/CA98/00607	(74) Agents:	ERRATT, Judy, A. et al.; Gowling, Strathy & Henderson, Suite 2600, 160 Elgin Street, Ottawa, Ontario K1P 1C3 (CA).
(22) International Filing Date:	25 June 1998 (25.06.98)	(81) Designated States:	AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, GW, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).
(30) Priority Data:	2,224,108 24 February 1998 (24.02.98)	CA	
(71) Applicant ( <i>for all designated States except US</i> ):	HER MAJESTY IN RIGHT OF CANADA as represented by THE MINISTER OF AGRICULTURE AND AGRI-FOOD CANADA [CA/CA]; Building 34, Centre for Food and Animal Research, Central Experimental Farm, Ottawa, Ontario K1A 0C6 (CA).		
(72) Inventors; and		Published	
(75) Inventors/Applicants ( <i>for US only</i> ):	GAVORA, Jan, S. [CA/CA]; 51 Forest Hill Avenue, Ottawa, Ontario K2C 1P7 (CA). FALCONER, Marcia, M. [CA/CA]; 33 Abingdon Drive, Nepean, Ontario K2H 7M5 (CA). NGUYEN, Thuy, H. [CA/CA]; Apartment 1201, 1316 Carling Avenue, Ottawa, Ontario K1Z 7L1 (CA). BENKEL, Bernhard, F. [CA/CA]; 293 Summit Avenue, Ottawa, Ontario K1H 5Z7 (CA).	With international search report.	

(54) Title: TRANS-SOMATICS WITH GENE TRANSFER INTO MAMMARY EPITHELIAL CELLS



(57) Abstract

A method is described to transfer a gene encoding a valuable compound, such as a pharmaceutical, into the secretory cells of the mammary gland to produce a new compound into the milk or to alter the composition of the milk. In this method the packaging cell line producing the viral-derived particles is infused into the mammary gland. The packaging cells will attach and survive for a period of time within the mammary gland. While the cells are viable, they will supply a continuous source of viral-derived particles to trans-infect the maximum number of mammary epithelial cells. After a period of time in the mammary gland, both the particles and the packaging cells will be destroyed by natural mechanisms while the trans-infected mammary epithelial cells continue to express gene(s) encoding the valuable compound or gene(s) to alter the composition of the milk. One or more genes can be trans-infected including DNA sequences that contribute to the efficient production of an active compound or to its stability. The packaging cells and the viral-derived particles used in this method can be those which trans-infect dividing or non-dividing cells and can be used either singly or together. This method can be used alone or in combination with other novel methods designed to ensure that the viral-derived particles are correctly positioned to trans-infect the mammary epithelial cells. Increasing the trans-infection of the mammary epithelial cells with the viral particles will result in a higher concentration of the valuable compound in the milk.

BEST AVAILABLE COPY

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	RS	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Boemia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NB	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

## TRANS-SOMATICS WITH GENE TRANSFER INTO MAMMARY EPITHELIAL CELLS

5

**FIELD OF INVENTION**

The present invention relates to a method of producing value-added milk by the incorporation of specific DNA sequence(s) into the cells of the mammary gland. The term "value-added milk" is meant to mean milk containing a valuable compound, 10 such as a pharmaceutical, as well as milk with a changed composition so that the market value of the milk is enhanced.

**BACKGROUND OF THE INVENTION**

This invention is based upon a technique to transfer a gene, and related non-translated control sequences, into the secretory cells of the mammary gland to produce 15 new compounds in milk and/or to alter the milk composition.

An example of a compound produced into milk by this method would be a pharmaceutical which cannot be manufactured in a biologically active state. This 20 invention provides an alternative method of producing such pharmaceuticals into milk using the synthetic capabilities of the mammary epithelial cell.

Mammals produce copious amounts of complex proteins into milk to provide nutrition for their young. To harness this capability, the DNA of, for example a 25 valuable pharmaceutical, can be inserted into a mammary epithelial cell and this cell then will produce the active, pharmaceutical compound into the secreted milk. In a cow, the udder will act as a receptacle to hold the milk until it can be collected. The milk containing the added compound can be processed to extract and purify this pharmaceutical compound for subsequent sale, possibly to the medical and/or 30 veterinary communities, or the milk could be consumed directly as a therapeutic

agent.

At present there are two basic techniques that can be used to produce value-added milk. The first is to create a transgenic animal by microinjection or transfection  
5 of foreign DNA into an ovum or a fertilized egg. Incorporation of DNA at this stage in development generally results in a transgenic animal which carries the inserted DNA in every cell.

There are positive and negative aspects of producing a transgenic animal  
10 which expresses foreign proteins in its milk. A positive point is that a single founder animal can create a population of transgenic animals by natural reproduction. However, progenies do not always produce the exogenous protein at the same level as the original animal. Negative aspects include the technically difficult procedures required to produce the animal and the long time between adding the foreign DNA  
15 and harvesting the exogenous protein. In addition, the presence of even very small amounts of an active pharmaceutical in every tissue may be detrimental to the health of the animal.

An alternative method is to produce value-added milk by adding the desired  
20 DNA only to the cells of the mammary gland of the animal. This results in a trans-somatic animal (or chimera) which contains the inserted DNA essentially in only one tissue, the mammary gland.

Trans-somatic animals have the advantage that they can be produced with less  
25 technical difficulty. They also can be produced quickly so that there is a period of only weeks to months between adding the foreign DNA and harvesting the exogenous protein. Moreover, since only one tissue contains the added DNA and produces the resulting compound, health risks to the trans-somatic animal are reduced. Although the DNA is not passed on to the progeny, this is compensated for by the ease and  
30 speed with which a trans-somatic animal can be produced.

A trans-somatic goat which expresses human growth hormone (hGH) into milk has been produced by Archer et al. (1994). In Archer the viral-derived particles were infused into the mammary gland for approximately every two days for two weeks. The levels of the compound, human growth hormone, which was used as an example, were very low and approached background levels after the first day. Also Gould et al. (United States Patent 5,215,904) described a method for increasing the rate of mitosis of mammary epithelial cells and then exposing these cells to viral particles for integration of the desired DNA into the epithelial cell.

10 To produce either a trans-somatic or a transgenic animal, exogenous DNA must pass through the exterior cell membrane. Eukaryotic cells have evolved a membrane which is impervious to most substances including heterologous DNA. Numerous techniques have been developed to bypass this barrier. These include:

15                   electroporation,  
                       carrier lipids (liposomes, negative, positive or neutral charged vesicles),  
                       mechanical wounding of cells including microinjection, liquid or air-jet pressure and scrape loading,  
                       use of particles composed partially or wholly of viral proteins.

20                   Methods which have successfully produced trans-somatic animals include:

1) Arterial injection of DNA carried in liposomes (small lipid vesicles) was used to produce a trans-somatic mouse (Thierry et al. 1995). This technique can be adapted to deliver DNA to the mammary gland by injection into the major artery serving the mammary gland but circulation of the blood carrying the DNA can result in transfection of multiple tissues.

25                   2) Direct injection of the DNA into tissues was used successfully to add foreign DNA to muscle and other tissues (Furth et al. 1992). This technique can be adapted to inject virus-like particles, carrying the foreign DNA, directly into the tissue of the udder.

30                   3) Use of viral-derived particles carrying DNA coding for human growth

hormone (hGH) were infused through the teat canal, for example by Archer et al. (1994). This resulted in production of trans-somatic goats which expressed hGH into the milk.

5       Viruses reproduce within cells and therefore have evolved a technique to bypass the protective cell membrane to deliver the viral genome (DNA) into a host cell. To enter a cell, protein(s) of the outer viral shells first bind to receptors on the cell surface and then the virus is internalized.

10      The method used by Archer involves transfecting a cell line with DNA coding for various, but not all, proteins of a virus. This cell line, called a "packaging cell line", will produce empty virus shells which can bind to receptors on the host cell membrane. When heterologous DNA, coding for a pharmaceutical or other milk modification, is transfected into the packaging cell line, this DNA will be packaged  
15     into the viral-derived particle. When the viral-derived particle comes in contact with a milk-producing cell, the viral proteins of the shell ensure that the heterologous DNA is carried into the cell. Other viral proteins, associated with the particle, integrate the heterologous DNA into the genome of the host cell so that the protein encoded by the DNA can be expressed. In this method the viral-derived particles are used to  
20     introduce the heterologous DNA into the mammary gland.

The trans-somatic methods of the prior art offer advantages over the transgenic method; however the very low levels of foreign protein in the milk of the trans-somatic animal have limited the commercial success of these methods.

25

Thus the present invention is directed to methods of improving the yield of the foreign protein in the milk of a trans-somatic animal.

## SUMMARY OF THE INVENTION

30      The present invention relates to a method of producing valuable compounds into milk and/or changing the composition of milk so as to enhance its properties

and/or its marketability. More specifically the present invention relates to the addition of specific DNA sequences, including non-translated regulatory sequences, to the cells of the mammary gland and the subsequent expression of compound(s) encoded by that DNA into the milk. Other DNA sequences that enhance the efficiency of production of the compound, enhance the stability of the compound, or result in biological activity of the compound can also be added to the mammary epithelial cells either at the same time or at a different time.

The novel method of the present invention involves the use of viral-derived particles and packaging cells which produce these particles for infusion into the mammary gland through the teat canal. The packaging cells will attach and survive for a period of time within the mammary gland. While the cells are viable, they will supply a continuous source of viral-derived particles. These viral-derived particles trans-infect only dividing cells and are destroyed relatively quickly in the mammary gland. Thus a continuous supply of viral-derived particles from the packaging cells present in the mammary gland will ensure that viral-derived particles are present and can trans-infect the mammary epithelial cells whenever they divide.

The packaging cells and viral-derived particles can be from retroviruses and from non-retroviruses. Most retroviral-derived particles trans-infect only dividing cells. Non-retroviral particles such as those from adenovirus, Epstein-Barr virus, or other viruses trans-infect non-dividing cells. Thus a mixture of the two types of particles, and the associated packaging cells if needed, will ensure delivery of the packaged DNA to the maximum number of cells. Moreover, DNA can be packaged into viral-derived particles *in vitro* and these can be used for trans-infection as well.

This method can be used alone or in combination with other novel methods designed to ensure that the viral-derived particles are correctly positioned to trans-infect the mammary epithelial cells. Increasing the trans-infection of the mammary epithelial cells with the viral particles will result in a higher concentration of the valuable compound in the milk along with possible other compounds produced from

the inserted DNA that will enhance the efficient production, stability or activity of the compound.

- Thus according to the present invention there is provided a method of  
5 producing a trans-somatic mammal, wherein said method provides the incorporation  
of a DNA sequence into the secretory cells of the mammary gland to alter the  
composition of the milk, comprising the steps of: providing a vector containing a  
DNA sequence encoding a valuable compound; packaging said vector into a cell line;  
preparing a solution comprising the packaged vector and cell line producing said  
10 packaged vector; and delivering said solution into the mammary gland to allow the  
incorporation of the DNA into the secretory cells of the mammary gland.

#### BRIEF DESCRIPTION OF THE DRAWINGS

- These and other features of the invention will become more apparent from the  
15 following description in which reference is made to the appended drawings wherein:  
**FIGURE 1** shows the presence of amylase in an udder infused with Clone 10 (left  
hind quarter) and Clone 12 (right hind quarter). The left front quarter was left  
untouched as a negative control and the right front quarter was infused with DMEM  
and Polybrene but no cells, to serve as a further negative control. Lanes 1, 2 & 3, early  
20 premilk from 3 quarters of cow #99. Lane 1, control quarter (RF-); Lane 2, treated  
quarter (RH+); Lane 3, treated quarter (LH+). Lanes 4, 5 & 6, late premilk from cow  
#99. Lane 4, treated quarter (LH+); Lane 5, treated quarter (RH+); Lane 6, control  
quarter (RH-). Lane 7, mixture of pig and chicken amylase standards; Lane 8, chicken  
amylase standard.  
25 **FIGURE 2** is a Western blot showing the presence of tPA in an udder infused with  
Clone 1. Lane 1, premilk from treated quarter (LH+) of cow #56; Lane 2, premilk  
from control quarter (LF-) of cow #56; Lane 3, premilk from treated quarter (RF+) of  
cow #90.

#### 30 DESCRIPTION OF THE PREFERRED EMBODIMENT

The present invention relates to a method to transfer a gene or genes, and

related non-translated control sequences into the secretory cells of a mammary gland to produce "value-added milk". The term "value-added milk" is meant to mean milk containing a valuable compound, such as a pharmaceutical, as well as milk with a changed composition so that the market value of the milk is enhanced.

5

Pharmaceuticals according to the present invention can include, but are not limited to: a tissue plasminogen activator, a blood clotting factor, an antibody, a protein to aid in weight reduction, a galactosyltransferase, a growth factor, an oncoprotein, a protease inhibitor, a hormone, a milk protein, a hormone receptor, a tumor suppressor protein, an aging inhibitor, or an erythropoietin.

10

According to the present invention there is provided a suitable DNA vector (plasmid) containing the desired DNA sequence encoding the heterologous protein (valuable compound) and associated regulatory sequences such as promoters, enhancers, introns, signal sequences, etc. Other sequences to produce compounds that increase efficient production, enhanced stability or are involved in biological activity of the compound can be included in the same or another DNA vector.

15

In one embodiment the vector is based on the pLXSN, pLNCX, or pLNSX plasmids provided under license from Fred Hutchinson Cancer Research Center.

20

According to the present invention the delivery of the DNA to the cells of the mammary gland is accomplished by the use of viral-derived particles and packaging cells, which produce these particles, for infusion into the mammary gland through the teat canal. This infusion is a standard veterinary practice and usually involves the use of a cannula for insertion into the teat canal, the cannula being attached to a syringe containing a solution of the packaging cells and viral-derived particles.

25

It has already been established that such particles derived from retrovirus trans-infect only dividing cells. Thus, the infusion of the particles and the packaging cells takes place when mammary cell division is naturally at a high level. In the

present invention heifers at 5 to 7 months pregnancy were used. It would also be possible to use non-pregnant mammals treated with hormones to induce mammary cell division and lactation prior to particle and cell infusion. However, this method would involve additional handling of the animal and thus is not preferred.

5

Other types of viral-derived particles and associated packaging cell lines, such as those based on adenovirus, Epstein-Barr virus, or other viruses can also be used.

These non-retroviral derived particles can infect non-dividing cells. If used in combination with the retroviral derived particles described above, the maximum

10 number of cells will receive the desired DNA. In addition, particles to deliver DNA can be manufactured *in vitro* and use of these alone or in combination with the above described particles will enhance the number of mammary epithelial cells which incorporate the desired DNA. An example of particles which are constructed *in vitro* is described or referred to in Morsy and Caskey (1997).

15

The present invention is applicable to all mammals and is especially applicable to all non-human mammals. Goats, sheep and cows are preferred. Cows with their inherent large volume milk production are particularly preferred.

20

The packaging cells will attach and survive for a period of time within the mammary gland. The reason for the presence of the cells is to supply a continuous source of the viral-derived particles to trans-infect the maximum number of mammary epithelial cells. Both the packaging cells and the viral-derived particles are eventually destroyed in the mammary gland; however, considering the fragile nature of the

25

particles *in vitro*, they are likely destroyed within days. The cells, by contrast, may persist much longer. When cells were infused into the udder, and the udder subsequently flushed and assayed for the presence of cells (dead and alive), most of the cells could not be flushed out. Cells were seen in the wash for 3 days after infusion; of those cells that were flushed out, a large majority remain viable,

30

suggesting that most of the cells have become attached to the udder. In prior art methods (for example Archer et al. 1994), the particles are infused into the udder

approximately every two days for two weeks. In the present invention no subsequent infusions are necessary. Thus an advantage of the present invention over the prior art is a reduction in the handling of the animal.

5        In prior arts methods (for example, Archer et al. 1994), large scale tissue culture was required to supply sufficient numbers of particles for multiple infusions. An advantage of the present invention over the prior art is that preparation of a large number of particles is not required.

10      In prior arts methods (for example, Archer et al. 1994), the preparation of a stock solution of viral-particles required ultra-centrifugation and resuspension of the particles. A further advantage of the present invention is that no ultracentrifugation is involved. Centrifugation and other handling techniques can destroy the relatively fragile particles. In the present invention, handling of the particles is minimal, 15     increasing the probability that the particles present will remain intact and useful.

The above cited advantages are in addition to the improvement in yield of the compound in the milk, which has already been described above.

20      This method can be used alone or in combination with other methods of the present invention to increase the incorporation of DNA from the viral-derived particles. The additional methods of the present invention are all designed to ensure that the viral-derived particles are correctly positioned to trans-infect the mammary epithelial cells.

25      In one embodiment, the teat canal and udder is first emptied by milking or under negative pressure. Then the udder is flushed with an osmotically-balanced solution which is infused into the teat canal and udder until the udder is full. This solution is removed by milking or under negative pressure and the desired suspension 30     is infused into the udder. In an alternative procedure, the teat canal is not emptied first but is directly filled with an osmotically-balanced solution which is then removed

10

by milking or under negative pressure.

Flushing of the udder serves two purposes; it removes the thick, secreted fluid that is normally present and it forces open the ductwork to allow better access of the  
5 packaging cells and viral-derived particles which will be inserted in the next step of the procedure. An example of a suitable osmotically-balanced solution includes saline solution, but can include any other buffered solutions, and also can include the packaging cell grown medium.

10 External massage applied several times a day to the mammary gland improves the circulation of the viral-derived particles and results in more secretory cells in the mammary gland being exposed to the viral-derived particles and thus more cells will incorporate the DNA carried in these particles.

15 In another embodiment, the packaging cells are grown on a commercially available matrix (designed to support growth and replication of tissue culture cells). This solid support matrix can include gelatin, glass, collagen or plastic beads. Cytodex beads or Cultisphere (purchased from Sigma) are two specific examples of useful support means. The beads, with cells adhering to them, are infused into the  
20 mammary gland through the teat canal along with a suspension of DNA-containing viral-derived particles. The beads remain in the mammary gland. External massage applied several times a day to the mammary gland recirculates the beads and helps to distribute the viral-derived particles that are produced by the packaging cells growing on the beads. The presence of the packaging cells on beads also ensures the  
25 continuous presence of many more DNA-containing viral-derived particles. Concomitantly more secretory cells in the mammary gland will incorporate the DNA carried within these viral particles and the production of the desired protein will increase accordingly. The beads and any cells remaining on them will eventually be removed when milking is begun.

30

In another embodiment, the infusion of cells (with or without growth on

beads) and viral-derived particles containing the desired DNA is followed by infusion of a substance which is more dense than the water-based suspension fluids, used for cell infusion, such as growth medium (Dulbecco's modified Eagle's medium [DMEM], phosphate buffered saline [PBS], etc.). This dense fluid, which in one example is composed of silicone, displaces the aqueous solutions containing the packaging cells and the viral-derived particles upwards into the ductwork of the mammary gland. This prevents collection of the packaging cells in the cistern of the udder and positions the viral-derived cells and the viral-derived particles up into the ducts of the mammary gland. Thus, the viral-derived particles are positioned near the dividing cells in the alveoli of the mammary gland and allow more of the DNA carried in the viral-derived particles to be incorporated. Any physiologically compatible inert fluid that has a density greater than that of the infusion solution can be used according to this embodiment of the present invention.

15        Thus, the present invention consists of the following procedure, which can be used alone or in combination with optional methods of the present invention, to deliver heterologous DNA to milk producing cells. In this procedure the following steps occur:

- 20        1) A vector containing the desired DNA sequence(s) is constructed and is transfected by standard means into a packaging cell line.
- 25        2) A solution containing the packaging cell line, producing viral-derived particles containing the desired DNA sequence(s), and viral-derived particles, is infused into the mammary gland through the teat canal. The packaging cells attach to the epithelial cells of the mammary gland, remain viable and produce viral-derived particles.
- 30        3) The DNA becomes incorporated into the secretory cells of the mammary gland.
- 4) The milk containing the product induced by addition of the desired DNA is milked from the cistern. The product is purified from the milk or the milk, containing the product, is consumed.

The basic technique can be modified by the addition of one or more of the following steps:

- 1) The udder is flushed with an osmotically-balanced solution.
- 2) The packaging cells are grown on a matrix designed to support proliferation of eukaryotic cells, such as Cytodex beads, and then infused through the teat canal. The cells will remain in the mammary gland and continue to produce viral-derived particles for a period of at least 3 days.
- 3) After infusion of the packaging cells (either as a suspension or grown on a matrix), a compound which is denser than aqueous solutions, such as silicone, is infused into the mammary gland to force the aqueous solutions carrying the packaging cells and viral-derived particles up and into the region of the mammary gland where the DNA can be incorporated into the milk producing cells. The dense compound, if it has been added, and any unattached cells as well as the matrix, if used, is removed from the mammary gland cistern by milking at an appropriate time after infusion. An appropriate time is defined as after the majority of packaging cells attach and after there is sufficient production of viral-derived particles. An example of an appropriate time would be at least 3 days, however shorter or longer periods may also be used. The packaging cells which are not removed at this time die and are removed by the recipient's natural mechanisms.
- 4) External massage applied several times a day to the mammary gland will recirculate the viral-derived particles and the packaging cells, provided either as a suspension or grown on a matrix. This helps to distribute the viral-particles and the cells and increases the incorporation of the DNA carried in these particles into the secretory cells.

25

While this invention is described in detail with particular reference to preferred embodiments thereof, said embodiments are offered to illustrate but do not limit the invention.

**EXAMPLES****Example 1: Preparation of plasmids carrying the desired gene and various control sequences for use in retroviral particle delivery to bovine mammary epithelial cells**

5 As a system to transfect the mammary epithelial cells with a desired gene, the PG13 packaging cell line was acquired under license from Fred Hutchinson Cancer Research Center. This system was used because it produces retroviral particles containing the gibbon ape leukemia virus envelope (Galv) which facilitates transfection of bovine cells. To package the desired gene "X" into retroviral particles, 10 the plasmids, pLXSN, pLNCX, pLNSX and pLN, also were acquired under license from Fred Hutchinson Cancer Research Center.

**Modifications of plasmids:**

In order to have an alternative method to select cell clones containing the 15 desired gene, we replaced the *neomycin (neo)* gene with the *hygromycin (hyg)* gene in pLXSN, pLNSX and pLNCX. This was done by long-range PCR-amplification of the region around the *neo* gene then ligating the PCR product with the *hyg* gene which was amplified from the plasmid pREP4 (purchased from Invitrogen). The resulting 20 plasmids were called pLXSH, pLHSX and pLHCX where "L" represents the Moloney murine virus long terminal repeat (LTR) acting as a promoter, "S" is the SV40 promoter and "C" is the cytomegalovirus promoter region.

To put the desired gene, "X", under the control of a constitutive promoter, the 25 cytomegalovirus promoter region in pLNCX and pLHCX was removed by restriction digestion and replaced with the beta actin promoter to make pLNAX and pLHAX. The beta actin promoter sequence was derived using PCR from the pJ6Ω plasmid purchased from ATCC (catalog no. 37723).

In order to avoid transcription interference between the gene used for selection 30 and the gene used to produce the desired protein, transcription from both genes was coupled by replacing the SV40 promoter from the plasmids pLXSN and pLXSH with

an internal ribosomal entry site (IRES). The resulting plasmids, pLAI<sub>N</sub> and pLAI<sub>H</sub>, express both the selection gene and the inserted gene under the same (LTR) promoter. The IRES in these plasmids provides the translation initiation site within the transcript, allowing the downstream gene product to be produced. The IRES used here is identical to the IRES found in the plasmid pIRES1neo purchased from Clontech. The sequences for pLAI<sub>N</sub> and pLAI<sub>H</sub>, where A in this case is chicken amylase but which can be any desired protein, are shown in SEQ ID No:1 and SEQ ID No: 2, respectively.

In order to increase retroviral titre or stability of the transcript, or to increase the expression level of the desired gene during lactation, or to allow translation of more than one protein from the same transcript, the basic plasmid, pLNCX, was modified. In one modification, the selection gene was removed to minimize the size of the resulting plasmid, pLX. In another modification, the CMV promoter was replaced with the murine mammary tumour virus (M) LTR promoter to improve transcription of the resulting plasmid, pLNMX, during lactation. The sequence for pLNMX is shown in SEQ ID No: 3.

In another modification, a wild-type IRES was modified so that the ATG codon at position 10 is destroyed and the sequence downstream of the ATG codon at position 11 codes for the desired gene, "X", in a plasmid such as pLN*M*<sub>i</sub>X. The sequence for this IRES modification (*i*<sub>2</sub>) is included within the sequence for pLN*M*<sub>i</sub>X, shown in SEQ ID No: 4.

**Example 2: Preparation of cell clones producing viral-derived particles carrying the chicken amylase gene as a "marker" protein**

The two packaging cell lines used in this experiment were purchased from ATCC, PA317 (catalog no. CRL-9078) and PG13 (catalog no. CRL-10686). A description of both PA317 and PG13 and their use can be found in Miller et al. 1990.

the retroviral particle delivery system, we used a stable and readily detectable marker protein, chicken amylase, in some experiments. Chicken amylase migrates at a unique position by native gel electrophoresis and can be differentiated from bovine or other amylases. To produce a PG13-derived cell line (a clone) producing retroviral particles which carried chicken amylase under the control of the beta actin promoter,  
5 the following procedure was done.

The pLH(A)amy plasmid with *hygromycin* (H) driven by the LTR promoter and with *amylase* (amy) driven by the beta actin promoter (A), was produced by  
10 standard recombinant techniques. pLH(A)amy was transiently transfected by the calcium phosphate technique into the PA317 packaging cell line. The transfected PA317 cells produce viral particles containing the amy RNA into the supernatant. The amphotropic viral envelope protein of these particles allows entry into cells of most species including the PG13 packaging cells in a process called trans-infection.  
15 Empirically it has been determined that trans-infection produces PG13 clones with a higher rate of particle production as compared to PG13 clones produced by other means of DNA insertion (particle bombardment, calcium phosphate or liposome transfection). Therefore, the viral-derived particles in the supernatant of the  
20 transfected PA317 cells were used to trans-infect the PG13 packaging cell line. The resulting clones were selected using 700 ug/ml hygromycin in DMEM with 10% fetal bovine serum (FBS) and 1% penicillin/streptomycin for 14 days. Clones were picked using cloning rings and grown as separate cell lines.

These PG13 amy clones were tested for amylase production by native gel  
25 electrophoresis of  $\alpha$ -amylase. The  $\alpha$ -amylase samples in loading buffer (0.1M tris-borate pH 8.5, 10% sucrose, 10mM magnesium chloride) were electrophoresed for 3-4 hours at 250 volts, with cooling, on a 5.5% acrylamide gel containing 0.1M tris-borate pH 8.5. The electrophoresis buffer was 0.1M tris-borate pH 8.5. The gel was then placed with gentle shaking for 1 h at room temperature in a solution of 2%  
30 soluble starch, 10mM calcium chloride, 50mM tris-HCl pH 7.5. The gel was briefly rinsed in water, then placed in a dilute solution of iodine ( $I_2 : KI : \text{water}$  is about 1 : 2 :

2000) until the amylase signal shows as a clear band on the gel.

All packaging cells produce filled retroviral particles which contain the desired gene and empty particles which do not carry the desired gene. To increase the number  
5 of filled retroviral particles, hygromycin-resistant PG13 amy clones producing high amounts of amylase were supertransfected with a second plasmid, pLN(A)amy, which also carried the amylase gene but which has *neomycin* as the clone selection agent. In this procedure, first pLN(A)amy was transfected into the PA317 cell line by the calcium phosphate procedure, and then the viral-derived particles in the  
10 supernatant of the transfected PA317 amy-containing cells were used to trans-infect a PG13/pLH(A)amy clone. The PG13/pLH(A) clone which was supertransfected was chosen on the basis of high amylase production. Superclones, containing both pLN(A)amy and pLH(A)amy sequences were obtained by selecting in the presence of 1000 ug/ml G418 and 700 ug/ml hygromycin as described above. The surviving  
15 clones were picked and grown up for further testing. Although a marker protein, amylase, was used in these experiments to optimize filling of particles, it is obvious that the same procedure can be done where another protein, such as a pharmaceutical protein, is used instead of the amylase.

20 The superclones producing the highest levels of amylase were analyzed for particle production by the following method. For each clone to be tested, the supernatant containing viral-derived particles was used to trans-infect HeLa 229 (human cervical carcinoma) (purchased from ATCC catalog no. CCL-2.1) and Et2 (bovine mammary) cells (provided by Dr. Boris Zavizion, University of Vermont, Burlington, Vt.). Depending on the resistance gene contained in the plasmid, the trans-infected cells were treated with either 700 ug/ml hygromycin or 1000 ug/ml neomycin, or both, and the resulting colonies were counted. Each colony is the result  
25 of one infective (filled) particle. The clone producing the highest number of colonies was selected to be grown for subsequent infusion into the heifer udder. In the specific example shown in Figure 1, clone 10 produced the highest number of viral-derived particles and was designated as PG13/LH/LN(A)amy.  
30

Before a clone can be infused into the udder, it must be tested to determine that it does not produce replication-competent (called "helper") virus. Gene therapy is based on the assumption that the viral-derived particles can carry the desired DNA (i.e. cDNA for amylase or another protein) into the target cell but that once in the cell, 5 the virus cannot reproduce itself. There are two methods to check if helper virus is being produced by the trans-infected packaging cell line. These are described in detail in Cepko, 1992. The first method is to analyze the supernatant of the HeLa and Et2 cells for horizontal spread of hygromycin and/or neomycin resistance. The second method is to look for the presence of reverse transcriptase above background 10 levels in the supernatant of the trans-infected HeLa and Et2 cells. If trans-infection with the supernatant from HeLa and Et2 cells does not produce hygromycin and/or neomycin resistant colonies and if there is no reverse transcriptase above the level found in control cells, then it can be assumed that the particle producing clone, in this case PG13/LH/LN(A)amy, is not producing "helper" virus and therefore can be used 15 for infusion into the udder. We followed the procedures detailed in the Cepko reference and found neither colony nor reverse transcriptase production.

**Example 3: Infusion of clones producing retroviral-derived particles which carry a desired gene such as the amylase marker protein**

20 In one example, the cell clone PG13/LH/LN(A)amy, which expresses chicken amylase as a marker protein, was infused into the udder of a 7 month pregnant heifer, #99. Before infusion with the cell clone, each quarter of the udder was flushed with 250ml of saline solution at 37° C using the catheter portion of an angiocath G-18 catheter attached to a 140 cc Monoject syringe with a Luer lock attachment (both 25 purchased from CDMV, Saint-Hyacinthe, Quebec). The infusate was then milked out and measured to determine the volume of cell solution which could be administered to this particular quarter of this specific heifer.

30 Two superclones, designated as PG13/LH/LN(A)amy Clone 10 and PG13/LH/LN(A)amy Clone 12, were grown in standard tissue culture flasks until the total cell number (in an appropriate number of flasks) was  $>2 \times 10^8$  cells/clone. The

cells were trypsinized, counted and resuspended at  $1 \times 10^8$  cells in 250 ml DMEM without serum but containing 80 ug/ml of Polybrene to facilitate particle adsorption to the mammary epithelial cells.

5       The udder has four separate quarters. The left front quarter was left untouched as a negative control (LF-). The front right quarter was flushed with saline and infused with 250 ml DMEM plus 80 ug/ml Polybrene but with NO cells added and served as a negative control (RF-). The right hind quarter was infused with  $1 \times 10^8$  Clone 12 cells in 250 ml DMEM plus Polybrene (RH+) and the left hind quarter was  
10      infused  $1 \times 10^8$  Clone 10 cells in 250 ml DMEM plus Polybrene (LH+). The infusate was not removed. The udder was massaged 3 times a day, morning, noon and night, to help to distribute the cells which tend to settle into the cistern of the udder. Three weeks post infusion, the udder was "stripped" (all secretions were milked by hand from each quarter) which removed about 7 to 10 ml per quarter of a viscous, serum-  
15      like fluid. This was designated as "early premilk". Eight weeks post infusion the udder was stripped again, producing about 25 ml per quarter of a slightly cloudy, viscous fluid designated as "late premilk". The samples from the right front control quarter RF(-) and the two infused quarters, the right hind, RH(+) and the left hind, LH(+) were analyzed for the presence of active amylase as previously described (see  
20      Figure 1).

Figure 1 shows that in the early premilk sample, amylase is present in both infused quarters (RH) and (LH) while the control quarter (RF) contains no amylase. In the late premilk, the negative control quarter (RF) continues to show no amylase activity. Clone 10, left hind quarter, shows the highest amylase activity. Clone 12, right hind quarter, has decreased activity compared to the early premilk sample, but a small amount of amylase activity was detected. The samples loaded represent equal volumes (100 ul/lane) of the premilk samples. No attempt was made to load equal amounts of protein. The apparent reduction in the amount of amylase present between  
25      early and late premilk samples is caused by increased volume of the late premilk samples relative to the volume of the early premilk samples.  
30

**Example 4: Preparation of a clone producing viral-derived particles that carry the tPA gene and infusion of this clone into the udder of a pregnant heifer**

Tissue plasminogen activator (tPA) is used to treat heart attack and stroke by dissolving blood clots. We selected human tPA to be the first pharmaceutical product to be produced by our method in bovine milk. The ptPA-K plasmid, containing the mutated cDNA sequence for human tPA (where the amino acids KHRR 296-299 was mutated to AAAA, i.e. the "K" mutation), was purchased from ATCC (American Type Culture Collection, catalog no. 68059). Two additional mutations were incorporated into ptPA-K to form ptPA-TNK: the amino acid threonine at position 103 was mutated to asparagine (the "T" mutation), and the amino acid asparagine at position 117 was mutated to glutamine (the "N" mutation). The tPA amino acid sequence and a description of the above modifications can be found in Pennica et al. 1983. Both mutations were produced by using mismatched oligonucleotides containing the altered nucleotide sequence as primers for PCR amplification. The tPA-TNK gene was subsequently excised by restriction digestion and ligated into the pLXSH plasmid to make pL(tPA)SH.

The resulting PG13(tPA) clones were analyzed for tPA production by colorimetric determination using Spectrozyme (# 444 purchased from American Diagnostica Inc.). Clones which showed high levels of tPA were then checked for production of filled particles by colony counts (as described previously) and were safety checked to insure that no replication competent virus was being produced and that there was no reverse transcriptase production above that found as background in HeLa or ET2 cells. None of the clones tested produced replication competent virus by either of these tests. The clones with the highest particle production and highest level of tPA production were selected to be grown for infusion into the udder. In the following example, clone PG13/L(tPA)SH-1, was chosen for infusion into the udder.

In one example, Clone 1 (PG13/L(tPA)SH-1), which produces tPA, was infused into the udder of a seven and a half month pregnant heifer, #90 and into the

udder of a six and a half month pregnant heifer, #56. Before infusion of the clone and particles produced by it, each quarter of the udder in both heifers was flushed with 250ml of saline as described in Example 3. After milking out the saline, the right rear (RR) quarter of heifer #90 received  $2.5 \times 10^7$  cells of Clone 1 plus 80 ug/ml Polybrene 5 in 250 ml of DMEM without serum while the right front (RF) quarter received  $1 \times 10^8$  cells of Clone 1 plus 80 ug/ml Polybrene in 250 ml of DMEM without serum. The left front (LF) quarter received 250 ml of DMEM as a control while the left rear (LR) quarter was the untreated control. Heifer # 56 received  $2.5 \times 10^7$  cells of Clone 1 plus 80 ug/ml Polybrene in 250 ml of DMEM without serum in the left front quarter (LF+) 10 and  $1 \times 10^8$  cells of Clone 1 plus 80 ug/ml Polybrene in 250 ml of DMEM without serum in the right rear quarter (RR+). The right front quarter received 250 ml of DMEM as a control (RF-) while the left front (LF-) quarter was the untreated control. In both heifers, the infusate was not removed and the udder was massaged three times daily to help distribute the cells and particles up into the ductwork.

15

Four weeks post-infusion, the four quarters of both heifers were stripped to remove 5 to 9ml of the viscous, serum-like premilk fluid. The premilk samples from treated and control quarters of both heifers were analyzed for the presence of tPA by Western blotting. A Western blot showing the tPA results obtained from the best 20 quarter of each heifer is shown in Figure 2.

For Western blotting, the premilk samples were adjusted to pH 4.5 with acetic acid and centrifuged at 13,000 x g to pellet the "curd" fraction. The "whey" fraction contained in the supernatant was collected and diluted 1:5 with sample buffer. 25 Samples of 20 ul were loaded onto a 7.5% SDS PAGE gel. To detect tPA, the gel was transferred to nitrocellulose and then blocked overnight in 5% bovine serum albumin (BSA). The blot was incubated for 2 hours with a polyclonal antibody to tPA, #385R, purchased from American Diagnostica, diluted 1:500 with PBS followed by extensive washing in PBS. The secondary antibody, horseradish peroxidase goat-30 anti-rabbit, diluted 1:5000 in PBS, was incubated with the blot for one hour then extensively washed. Detection of antibody staining of tPA was by enhanced chemi-

luminescence (ECL) Amersham Detect Kit. Specificity of the primary antibody was previously determined by Western blotting a control sample of commercial tPA purchased from American Diagnostica.

5           Results of Western blotting samples obtained from the untreated control  
quarter (LF-) and from one treated quarters of each heifer are shown in Figure 2.  
Lane 1 shows a strong tPA band in the premilk from the treated (LH+) quarter of  
heifer #56. The middle lane, lane 2, contains premilk from the (LF-) control quarter  
of cow #56. Lane 3 shows a tPA band in the premilk from the treated (RF+) quarter  
10          of heifer #90. A faint smear in all 3 sample lanes is the result of non-specific binding  
of the antibody to an unidentified protein and should be disregarded.

{  
15          **Example 5: Determination of viability and attachment of 3T3 cells to the surface  
of the bovine udder.**

The cells used in this experiment were 3T3 cells (purchased from ATCC,  
catalog no. CCL-92) which had been transfected with luciferase as a marker protein.  
These 3T3 cells are essentially identical to the PG13 packaging cells except that the  
3T3 cells have not been transfected with the retroviral genes that are found in the  
20          packaging line.

Three quarters of the udder of a 5 month pregnant heifer and three quarters of  
the udder of a 7 month pregnant heifer were flushed with saline solution at 37° C  
using a standard veterinary infusion apparatus. The saline was then milked out and  
measured to determine what quantity of medium with or without cells could be  
25          infused into that particular quarter. In both heifers, one udder quarter was left  
untreated as a negative control; a second quarter received only DMEM, the third  
quarter received 1 X 10<sup>6</sup> cells suspended in an appropriate amount of DMEM, and the  
fourth quarter received 1 X 10<sup>7</sup> cells also suspended in an appropriate amount of  
30          DMEM. The quarters were then milked to produce about 10 ml of fluid at 3 hours, 24  
hours, and 72 hours post-infusion.

The fluid obtained by milking was divided in half. One portion was cultured under standard conditions for 3T3 cells to determine if the cells were viable. The second portion was used in a luciferase assay to count the approximate number of 3T3-luciferase cells per ml in the fluid. A control was run simultaneously using known numbers of 3T3-luciferase cells and a curve drawn to determine the correlation of luciferase intensity with cell number.

The results of two experiments indicated that the highest number of cells were found 3 hours post infusion and progressively fewer cells were found in the fluid milked from the udder at 24 and 72 hours. However the cells which were milked out at 72 hours were viable. This is evidence that the cells are attaching to the udder surface and that those cells which were unattached (and therefore could be milked out of the udder) remained viable for the period of the experiment. This suggests that the 3T3 cells are not quickly destroyed by the factors in the environment of the udder.

Thus PG13(tPA) cells can be expected to remain viable for at least a period of 3 days during which time they will continue to produce viral-derived particles and can therefore trans-infect a substantial number of mammary epithelial cells.

**Table 1: Attachment of 3T3 cells to udder walls**

Udder Quarter	Treatment	Attachment of 3T3 cells to udder interior at:		
		3 hours post-infusion	24 hours post-infusion	72 hours post-infusion
left hind	no treatment	-	-	-
left front	medium only	-	-	-
right hind	1 X 10 <sup>6</sup> cells	-	+/-	+
right front	1 x 10 <sup>7</sup> cells	+/-	++	++

**Example 6: Infusion of 3T3 cells grown on a solid support means**

The cells used in this experiment were 3T3 cells, as described in Example 2.

In this example however the cells were grown on Cytodex beads.

The cells growing on beads were infused into the udder of a 7 month pregnant heifer and subsequently removed 1, 2 and 3 days later as described in Example 2. Cells  
5 were removed from beads and viability was determined by Trypan blue exclusion. The cells remained viable for the entire period.

**Example 7: Use of Silicone to displace the cells and viral-derived particles**

The cells used in this experiment were 3T3 cells, as described in Example 3.  
10 The volume of the solution containing the cells and viral-derived particles was reduced by 50 ml, 25 ml, or 10 ml but the overall number of the cells remained constant.

15 After the solution was infused into the udder, as described in the preceding examples, approximately 50 ml, 25 ml or 10 ml of silicone was infused into the mammary gland, using the methods previously described.

20 After three days the silicone was removed from the cistern by milking. No reaction to the silicone was noted.

25 All scientific publications and patent documents are incorporated herein by reference.

**References:**

Archer, J.S., W.S. Kennan, M.N. Gould, R.D. Bremel. 1994. Human growth hormone (hGH) secretion in milk of goats after direct transfer of the hGH gene into the mammary gland by using replication-defective retrovirus vectors. Proc. Natl. Acad. Sci. USA, 91:6840-6844.

30 Cepko, C. 1992. Transduction of genes using retrovirus vectors. In: Short Protocols

in Molecular Biology, 2nd edition. Ed. F.M. Ausubel, R. Brent, R.E. Kingston, D.D. Moore, J.G. Seidman, J.A. Struhl. Pub. by John Wiley & Sons. New York, N.Y. pp 9-30 to 9-45.

5 Furth, P.A., A. Shamay, R.J. Wall, L. Henninghausen. 1992. Gene transfer into somatic tissues by jet injection. *Analyt. Biochem.* 205:365-368.

Miller, A.D., D.G. Miller, J. V. Garcia, C.M. Lynch. 1990. Use of retroviral vectors for gene transfer and expression. *Meth. Enzymol.* 217: 581-599.

10 Morsy, M.A. and C.T. Caskey. 1997. Safe gene vectors made simpler. *Nature Biotech.* 15:17.

15 Pennica, D, W.E. Holmes, W.J. Kohr, R.N. Harkins, G.A. Vehaar, C.A. Ward, W.F. Bennett, E. Yelverton, P.H. Seeburg, H.L. Heyneker, D.V. Goeddel. 1983. Cloning and expression of human tissue-type plaminogen activator cDNA in *E. coli*. *Nature* 301: 214-221.

20 Thierry, A.R., Y. Lunardi-Iskandar, J.L. Bryant, P. Rabinovich, R.C. Gallo, L.C. Mahan. 1995. Systemic gene therapy: Biodistribution and long-term expression of a transgene in mice. *Proc. Natl. Acad. Sci. USA.* 92:9742-9746.

25 The present invention has been described with regard to preferred embodiments. However, it will be obvious to persons skilled in the art that a number of variations and modifications can be made without departing from the scope of the invention as described in the following claims.

## SEQUENCE LISTING

## (1) GENERAL INFORMATION:

## (i) APPLICANT:

- (A) NAME: Her Majesty in Right of Canada as Rep. by Agriculture and Agri-Food Canada
- (B) STREET: Experimental Farm
- (C) CITY: Ottawa
- (D) STATE: Ontario
- (E) COUNTRY: Canada
- (F) POSTAL CODE (ZIP): K1A 0C6

(ii) TITLE OF INVENTION: Production of Value-added Milk by Incorporation of Specific DNA Sequences into Mammary Epithelial Cells

(iii) NUMBER OF SEQUENCES: 4

## (iv) COMPUTER READABLE FORM:

- (A) MEDIUM TYPE: Floppy disk
- (B) COMPUTER: IBM PC compatible
- (C) OPERATING SYSTEM: PC-DOS/MS-DOS
- (D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

## (v) CURRENT APPLICATION DATA:

APPLICATION NUMBER: CA 0,000,000

## (vi) PRIOR APPLICATION DATA:

- (A) APPLICATION NUMBER: CA 2,199,212
- (B) FILING DATE: 05-MAR-1997

## (2) INFORMATION FOR SEQ ID NO: 1:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 7699 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: circular

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

GAATTGCTAG CAATTGCTAG CAATTGCTAG CAATTCTACAC CAGATCACCG AAAACTGTCC

60

TCCAAATGTG TCCCCCTCAC ACTCCCAAAT TCGCGGGCTT CTGCCTCTTA GACCACTCTA

120

CCCTTATTCCC CACACTCACC GGAGCCAAAG CGCGGGCCCT TCCGTTCTT TGCTTTGAA	180
AGACCCCCACC CGTAGGTGGC AAGCTAGCTT AAGTAACGCC ACTTTGCAAG GCATGGAAAA	240
ATACATAACT GAGAATAGAA AAGTCAGAT CAAGGTCAGG AACAAAGAAA CAGCTGAATA	300
CCAAACAGGA TATCTGTGGT AAGCGGTTCC TGCCCCGGCT CAGGGCCAAG AACAGATGAG	360
ACAGCTGAGT GATGGGCCAA ACAGGATATC TGTGGTAAGC AGTTCCGCC CGGGCTCGGG	420
GCCAAGAACCA GATGGTCCCC AGATGCGGTC CAGCCCTCAG CAGTTCTAG TGAATCATCA	480
GATGTTCCA GGGTGCCCCA AGGACCTGAA AATGACCTG TACCTTATTT GAACTAACCA	540
ATCAGTTCGC TTCTCGCTTC TGTCGCGCG CTTCCGCTCT CCGAGCTCAA TAAAAGAGCC	600
CACAACCCCT CACTCGCGC GCCAGTCITC CGATAGACTG CGTCGCCCG GTACCCGTAT	660
TCCCAATAAA GCCTCTTGCT GTTTGCATCC GAATCGTGGT CTCGCTGTTG CTTGGGAGGG	720
TCTCCTCTGA GTGATTGACT ACCCACGACG GGGGTCTTTC ATTTGGGGGC TCGTCCGGGA	780
TTTGGAGACC CCTGCCAGG GACCACCGAC CCACCACCGG GAGGTAAGCT GGCCAGCAAC	840
TTATCTGTGT CTGTCCGATT GTCTAGTGT TATGTTGAT GTTATGCGCC TCGCTCTGTA	900
CTAGTTAGCT AACTAGCTCT GTATCTGGCG GACCCGTGGT GGAACTGACG AGTTCTGAAC	960
ACCCGGCCGC AACCTGGGA GACGTCCAG GGACTTTGGG GGCGTCTTTT GTGGCCCGAC	1020
CTGAGGAAGG GAGTCGATGT GGAATCCGAC CCCGTCAGGA TATGTGGTTC TGGTAGGAGA	1080
CGAGAACCTA AAACAGTTCC CGCCTCCGTC TGAATTCTTG CTTTCGGTTT GGAACCGAAG	1140
CCGCGCGTCT TGTCTGCTGC AGCGCTGCAG CATCGTTCTG TGTTGTCTCT GTCTGACTGT	1200
GTTCCTGTAT TTGTCTGAAA ATTAGGGCCA GACTGTTACC ACTCCCTTAA GTTTGACCTT	1260
AGGTCACTGG AAAGATGTCG AGCGGATCGC TCACAACCAG TCGGTAGATG TCAAGAAGAG	1320
ACGTTGGTT ACCTTCTGCT CTGCAGAACGCCAACCTTT AACGTCGGAT GGCCCGGAGA	1380
CGGCACCTTT AACCGAGACC TCATCACCCA GGTTAAGATC AAGGTCTTTT CACCTGGCCC	1440
GCATGGACAC CCAGACCAGG TCCCCTACAT CGTGACCTGG GAAGCCTTGG CTTTGACCC	1500
CCCTCCCTGG GTCAAGCCCT TTGTACACCC TAAGCCTCCG CCTCCTCTTC CTCCATCCGC	1560
CCCGTCTCTC CCCCTTGAAAC CTCCTCGTTC GACCCCGCCT CGATCCTCCC TTTATCCAGC	1620
CCTCACTCCT TCTCTAGGCG CCGGAATTGCG TAAACTCGAC ATGGAAGTCC TTCTCCTCCT	1680
CGCAGCTGTC GGGCTTTGCT GGGCACAGTA CAATCCCAAC ACTCAGGCTG GGAGGACATC	1740
TATCGTGCAT CTCTTTGAAT GGCGCTGGGC CGACATTGCA CTGGAGTGC G AACACTATTT	1800

AGCTCCTAAT GGGTTGGAG GAGTCAGGT TTCTCCTCCA AATGAAAACA TTGTCATTAC	1860
TAATCCGAAC AGGCCCTGGT GGGAAAGATA CCAGCCCAC AGCTACAAGA TCTGCAGTCG	1920
ATCGGGCAAT GAAAATGAAT TCAGAGACAT GGTGACCAGA TGCAACAATG TTGGAGTTCG	1980
TATTTATGTG GATGCTGTTG TCAATCACAT GTGTGGATCT ATGGGTGGCA CGGGCACCCA	2040
CTCAACATGT GGGAGCTATT TCAACACCGG GACTAGAGAT TTTCCCGCTG TGCGTACTC	2100
TGCCTGGGAT TTCAATGACG GCAAATGTCA CACTGCAAGT GGAGACATCG AAAATTATGG	2160
GGACATGTAT CAGGTCCGGG ATTGCAAGTT GTCCAGCCTT CTTGATCTGG CTCTGGAGAA	2220
GGACTATGTA CGCTCAACAA TTGCAGCGTA CATGAATCAC CTCATTGATA TGGGTGTAGC	2280
AGGGTTCCGG ATCGATGCTG CCAAGCATAT GTGCCAGGG GACATAAGAG CGTTTCTGGA	2340
CAAAC TGAC GATCTAAATA CTCAGTGGTT TTCAGCAGGA ACGAAACCCCT TTATTTACCA	2400
AGAGGTAATT GACTGGGAG GAGAGCCAAT CACAGGCAGT CAGTACTTTG GGAATGGCCG	2460
CGTGACAGAA TTCAAGTATG GTGCCAAACT GGGGACGGTG ATCCGGAAGT GGAATGGAGA	2520
GAAGATGGCC TACTAAAGA ACTGGGGAGA AGGCTGGGC TTTGTGCCCTT CTGACAGAGC	2580
CCTGGTGTGTT GTGGATAACC ACGACAACCA GCGGGGGCAC GGGGCAGGCG GAGCTTCCAT	2640
TCTTACTTTG TGGGATGCCA GGCTTTATAA AATGGCGGTT GGTTTCATGC TCGCTCATCC	2700
GTACGGGTTG ACACGGGTGA TGTCAAGTTA TCGTTGGCCA AGATATTCG AAAACGGAGT	2760
GGATGTTAAC GACTGGGTGG GACCACCAAG TAACTCGGAC GGATCGACGA AGTCCGTTAC	2820
AATCAACGCA GACACTACCT GTGGCAATGA CTGGGTCTGC GAACATCGCT GGCGACAAAT	2880
AAGGAACATG GTTATCTTCC GTAATGTGGT AGACGGTCAG CCTTTCTCAA ACTGGTGGGA	2940
CAACGGGAGC AATCAAGTAG CTTCGGTG CCGCGACAGA GGCTTCATTG TCTTTAATAA	3000
TGATGACTGG TATATGAATG TCGATTGCA AACTGGTCTG CCTGCTGGAA CCTACTGCGA	3060
TGTTATTTCT GGACAAAAGG AAGGCAGTGC GTGTACTGGA AAGCAGGTGT ACGTTTCCTC	3120
GGATGGAAAG GCCAATTTCAG AGATTAGTAA CAGCGATGAA GATCCATTG TTGCAATTCA	3180
CGTTGATGCC AAGTTATAAG CTTCGAGGAT CCACTAGTAA CGGCCGCCAG TGTGCTGGAA	3240
TTCGGCTTGT CGACATCTAG GGCGCCAAT TCCGCCCTC TCCCCCCCCC CCCTAACGTT	3300
ACTGGCCGAA GCCGCTTGGAA ATAAGGCCGG TGTGTGTTG TCTATATGTG ATTTTCCACC	3360
ATATTGCCGT CTTTGGCAA TGTGAGGGCC CGGAAACCTG GCCCTGTCTT CTTGACGAGC	3420
ATTCCCTAGGG GTCTTTCCCC TCTGCCAAA GGAATGCAAG GTCTGTTGAA TGTGGTGAAG	3480

GAAGCAGTTC	CTCTGGAAGC	TTCTTGAAAGA	CAAACAAACGT	CTGTAGCGAC	CCTTTGCAGG	3540
CAGCGGAACC	CCCCCACCTGG	CGACAGGTGC	CTCTGCGGCC	AAAAGCCACG	TGTATAAGAT	3600
ACACCTGCAA	AGGCAGGCACA	ACCCCAGTGC	CACGTTGTGA	GTTGGATAGT	TGTGGAAAGA	3660
GTCAAATGGC	TCTCCTCAAG	CGTAGTCAAC	AAGGGGCTGA	AGGATGCCCA	GAAGGTACCC	3720
CATTGTATGG	GAATCTGATC	TGGGGCCTCG	GTGCACATGC	TTTACATGTG	TTTAGTCGAG	3780
GTTAAAAAAAG	CTCTAGGCC	CCCAGACCAC	GGGGACGTGG	TTTCCTTTG	AAAAACACGA	3840
TGATAAGCTT	GCCACAAACCC	GGGATAATT	CTGCAGCCAA	TATGGGATCG	GCCATTGAAC	3900
AAGATGGATT	GCACGCAGGT	TCTCCGGCCG	CTTGGGTGGA	GAGGCTATT	GGCTATGACT	3960
GGGCACAACA	GACAATCGGC	TGCTCTGATG	CCGCCGTGTT	CCGGCTGTCA	GCGCAGGGC	4020
GCCCCGTTCT	TTTTGTCAAG	ACCGACCTGT	CCGGTGCCT	GAATGAACTG	CAGGACGAGG	4080
CAGCGCGGCT	ATCGTGGCTG	GCCACGACGG	GCGTTCTTG	CGCAGCTGTG	CTCGACGTTG	4140
TCACTGAAGC	GGGAAGGGAC	TGGCTGCTAT	TGGGCGAAGT	GCCGGGGCAG	GATCTCCTGT	4200
CATCTCACCT	TGCTCCTGCC	GAGAAAGTAT	CCATCATGGC	TGATGCAATG	CGGCGGCTGC	4260
ATACGCTTGA	TCCGGCTACC	TGCCATTG	ACCACCAAGC	GAAACATCGC	ATCGAGCGAG	4320
CACGTACTCG	GATGGAAGCC	GGTCTTGTG	ATCAGGATGA	TCTGGACGAA	GAGCATCAGG	4380
GGCTCGCGCC	AGCCGAACTG	TTCGCCAGGC	TCAAGGCGCG	CATGCCGAC	GGCGATGATC	4440
TCGTCGTGAC	CCATGGCGAT	GCCTGCTTGC	CGAATATCAT	GGTGGAAAAT	GGCCGCTTT	4500
CTGGATTCAT	CGACTGTGGC	CGGCTGGGTG	TGGCGGACCG	CTATCAGGAC	ATAGCGTTGG	4560
CTACCCGTGA	TATTGCTGAA	GAGCTTGGCG	GCGAATGGGC	TGACCGCTTC	CTCGTGCTTT	4620
ACGGTATCGC	CGCTCCCGAT	TCGCAGCGCA	TCGCCTTCTA	TCGCCTTCTT	GACGAGTTCT	4680
TCTGAGCGGG	ACTCTGGGT	TCGATAAAAT	AAAAGATTTT	ATTTAGTCTC	CAGAAAAAGG	4740
GGGGAATGAA	AGACCCCACC	TGTAGGTTTG	GCAAGCTAGC	TTAAGTAACG	CCATTTGCA	4800
AGGCATGGAA	AAATACATAA	CTGAGAATAG	AGAAGTTCA	ATCAAGGTCA	GGAACAGATG	4860
GAACAGCTGA	ATATGGCCA	AACAGGATAT	CTGTGGTAAG	CAGTTCTGC	CCCGGCTCAG	4920
GGCCAAGAAC	AGATGGAACA	GCTGAATATG	GGCCAAACAG	GATATCTGTG	GTAAGCAGTT	4980
CCTGCCCGG	CTCAGGGCCA	AGAACAGATG	GTCCCCAGAT	GCGGTCCAGC	CCTCAGCAGT	5040
TTCTAGAGAA	CCATCAGATG	TTTCCAGGGT	GCCCCAAGGA	CCTGAAATGA	CCCTGTGCCT	5100
TATTTGAACT	AACCAATCAG	TTCGCTTCTC	GCTTCTGTTC	GCGCGCTTCT	GCTCCCCGAG	5160

CTCAATAAAA GAGCCCACAA CCCCTCACTC GGGGCGCCAG TCCTCCGATT GACTGAGTCG	5220
CCCGGGTACC CGTGTATCCA ATAAACCTC TTGCAGTTGC ATCCGACTTG TGGTCTCGCT	5280
GTTCCCTTGGG AGGGTCTCCT CTGAGTGAATT GACTACCGT CAGCGGGGGT CTTTCATTG	5340
GGGGCTCGTC CGGGATCGGG AGACCCCTGC CCAGGGACCA CCGACCCACC ACCGGGAGGT	5400
AAGCTGGCTG CCTCGCGCGT TTCGGTGATG ACGGTGAAAA CCTCTGACAC ATGCAGCTCC	5460
CGGAGACGGT CACAGCTTGT CTGTAAGCGG ATGCCGGGAG CAGACAAGCC CGTCAGGGCG	5520
CGTCAGCGGG TGTTGGCGGG TGTCGGGCG CAGCCATGAC CCAGTCACGT AGCGATAGCG	5580
GAGTGTATACT TGGCTTAACT ATGCCGCATC AGAGCAGATT GTACTGAGAG TGCACCATAT	5640
GCGGTGTGAA ATACCGACA GATCGTAAG GAGAAAATAC CGCATCAGGC GCTCTTCCGC	5700
TTCCTCGCTC ACTGACTCGC TGCCTCGGT CGTTCGGCTG CGCGAGCGG TATCAGCTCA	5760
CTCAAAGGCG GTAATACGGT TATCCACAGA ATCAGGGAT AACGCAGGAA AGAACATGTG	5820
AGCAAAAGGC CAGCAAAAGG CCAGGAACCG TAAAAAGGCC GCGTTGCTGG CGTTTTCCA	5880
TAGGCTCCGC CCCCTGACG AGCATCACAA AAATCGACGC TCAAGTCAGA GGTGGCGAAA	5940
CCCGACAGGA CTATAAAAGAT ACCAGGCCTT TCCCCCTGGA AGCTCCCTCG TGCGCTCTCC	6000
TGTTCCGACC CTGCCGCTTA CCGGATACCT GTCCGCCTTT CTCCCTCGG GAAGCGTGGC	6060
GCTTTCTCAT AGTCACGCT GTAGGTATCT CAGTTCGGTG TAGGTCGTTG GCTCCAAGCT	6120
GGGCTGTGTG CACGAACCCC CGGTTCAGCC CGACCGCTGC GCCTTATCCG GTAATATCG	6180
TCTTGAGTCC AACCCGGTAA GACACGACTT ATGCCACTG GCAGCAGCCA CTGGTAACAG	6240
GATTAGCAGA GCGAGGTATG TAGGCGGTGC TACAGAGTTC TTGAAGTGGT GGCCTAACTA	6300
CGGCTACACT AGAAGGACAG TATTGGTAT CTGCGCTCTG CTGAAGCCAG TTACCTTCGG	6360
AAAAAGAGTT GGTAGCTCTT GATCCGGCAA ACAAAACCACC GCTGGTAGCG GTGGTTTTTT	6420
TGTTTGCAAG CAGCAGATTA CGCGCAGAAA AAAAGGATCT CAAGAAGATC CTTTGATCTT	6480
TTCTACGGGG TCTGACGCTC AGTGGAACGA AAACTCACGT TAAGGGATTT TGGTCATGAG	6540
ATTATCAAAA AGGATCTCA CCTAGATCCT TTTAAATTAA AAATGAAGTT TTAAATCAAT	6600
CTAAAGTATA TATGAGTAAA CTTGGTCTGA CAGTTACCAA TGCTTAATCA GTGAGGCACC	6660
TATCTCAGCG ATCTGTCTAT TTCTGTCATC CATAGTTGCC TGACTCCCCG TCGTAGAT	6720
AACTACGATA CGGGAGGGCT TACCATCTGG CCCCAGTGCT GCAATGATAC CGCGAGACCC	6780
ACGCTCACCG GCTCCAGATT TATCAGCAAT AAACCAAGCCA GCCGGAAGGG CCGAGCGCAG	6840

AAGTGGTCCT GCAACTTTAT CCGCCTCCAT CCAGTCTATT AATTGTTGCC	GGGAAGCTAG	6900
AGTAAGTAGT TCGCCAGTTA ATAGTTGCG CAACGTTGTT GCCATTGCTG	CAGGCATCGT	6960
GGTGTACCGC TCGTCGTTTG GTATGGCTTC ATTCAAGCTCC	GGTTCCAAC GATCAAGGCG	7020
AGTTACATGA TCCCCATGT TGTGAAAAAA AGCGGTTAGC TCCTTCGGTC	CTCCGATCGT	7080
TGTCAGAAGT AAGTTGGCCG CAGTGTATC ACTCATGGTT	ATGGCAGCAC TGCATAATT	7140
TCTTACTGTC ATGCCATCCG TAAGATGCTT TTCTGTGACT	GGTGAGTACT CAACCAAGTC	7200
ATTCTGAGAA TAGTGTATGC GGCGACCGAG TTGCTCTTC	CCGGCGTCAA CACGGGATAA	7260
TACCGCGCCA CATAGCAGAA CTTTAAAAGT GCTCATCATT	GGAAAACGTT CTTCGGGGCG	7320
AAAACCTCTCA AGGATCTTAC CGCTGTTGAG ATCCAGTCG	ATGTAACCCA CTCGTGCACC	7380
CAA CTGATCT TCAGCATCTT TTACTTCAC CAGCGTTCT	GGGTGAGCAA AAACAGGAAG	7440
GCAAAATGCC GCAAAAAAGG GAATAAGGC GACACGGAAA	TGTTGAATAC TCATACTCTT	7500
CCTTTTCATA TATTATTGAA GCATTTATCA GGGTTATTGT	CTCATGAGCG GATACATATT	7560
TGAATGTATT TAGAAAAATA AACAAATAGG GGTTCCGCGC	ACATTTCCCC GAAAAGTGCC	7620
ACCTGACGTC TAAGAAACCA TTATTATCAT GACATTAACC	TATAAAAATA GGCGTATCAC	7680
GAGGCCCTTT CGTCTCAA		7699

## (2) INFORMATION FOR SEQ ID NO: 2:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 7980 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: circular

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

GAATTGCTAG CAATTGCTAG CAATTGCTAG CAATTCAAC CAGATCACCG	AAAACGTCC	60
TCCAAATGTG TCCCCCTCAC ACTCCCAAAT TCGCGGGCTT CTGCCTTTA	GACCACTCTA	120
CCCTATTCCC CACACTCACCG GGAGCCAAAG CCGCGGCCCT	TCCGTTTCTT TGCTTTGAA	180
AGACCCCAACC CGTAGGTGGC AAGCTAGCTT AAGTAACGCC	ACTTTGCAAG GCATGGAAA	240
ATACATAACT GAGAATAGAA AAGTCAGAT CAAGGTCAGG	AAACAAAGAAA CAGCTGAATA	300
CCAAACAGGA TATCTGTGGT AAGCGGTTCC	TGCCCCGGCT CAGGGCCAAG AACAGATGAG	360

ACAGCTGAGT GATGGGCCAA ACAGGATATC TGTGGTAAGC AGTTCCCGCC CCGGCTCGGG	420
GCCAAGAAC A GATGGTCCCC AGATGCGGTC CAGCCCTCAG CAGTTTCTAG TGAATCATCA	480
GATGTTCCA GGGTGCCCCA AGGACCTGAA AATGACCCTG TACCTTATTG GAACTAACCA	540
ATCAGTTCGC TTCTCGCTTC TGTTCGCGCG CTTCCGCTCT CCGAGCTCAA TAAAAGAGCC	600
CACAACCCCT CACTCGGCAGC GCCAGTCTTC CGATAGACTG CGTCGCCCGG GTACCCGTAT	660
TCCCAATAAA GCCTCTTGCT GTTTGCATCC GAATCGTGGT CTCGCTGTTG CTTGGGAGGG	720
TCTCCTCTGA GTGATTGACT ACCCAGACG GGGGTCTTTC ATTTGGGGC TCGTCCGGGA	780
TTTGGAGACC CCTGCCAGG GACCACCGAC CCACCACCGG GAGGTAAGCT GGCCAGCAAC	840
TTATCTGTGT CTGTCGATT GTCTAGTGTC TATGTTGAT GTTATGCGCC TGGTCTGTA	900
CTAGTTAGCT AACTAGCTCT GTATCTGGCG GACCCGTGGT GGAACGTGACG AGTTCTGAAC	960
ACCCGGCCGC AACCTGGGA GACGTCCCGAG GGACTTTGGG GGCGTTTTT GTGGCCCGAC	1020
CTGAGGAAGG GAGTCGATGT GGAATCCGAC CCCGTCAGGA TATGTGGTTC TGGTAGGAGA	1080
CGAGAACCTA AAACAGTTCC CGCCTCCGTC TGAATTTTG CTTTCGGTTT GGAACCGAAG	1140
CCGCGCGTCT TGTCTGCTGC AGCGCTGCAG CATCGTTCTG TGGTCTCT GTCTGACTGT	1200
GTTTCTGTAT TTGTCTGAAA ATTAGGGCCA GACTGTTACC ACTCCCTTAA GTTTGACCTT	1260
AGGTCACTGG AAAGATGTCG AGCGGATCGC TCACAACCAG TCGGTAGATG TCAAGAAGAG	1320
ACGTTGGGTT ACCTCTGCT CTGCAGAATG GCCAACCTTT AACGTCGGAT GGCGCGAGA	1380
CGGCACCTTT AACCGAGACC TCATCACCCCA GTTAAAGATC AAGGTCTTTT CACCTGGCCC	1440
GCATGGACAC CCAGACCAGG TCCCCATACAT CGTGACCTGG GAAGCCTTGG CTTTGACCC	1500
CCCTCCCTGG GTCAAGCCCT TTGTACACCC TAAGCCTCCG CCTCCTCTTC CTCCATCCGC	1560
CCCGTCTCTC CCCCTGAAC CTCCTCGTTC GACCCCGCCT CGATCCTCCC TTTATCCAGC	1620
CCTCACTCCT TCTCTAGGCG CCGGAATTGCG TAAACTCGAC ATGGAAGTCC TTCTCCTCCT	1680
CTCAGCTGTC GGGCTTGCT GGGCACAGTA CAATCCAAAC ACTCAGGCTG GGAGGACATC	1740
TATCGTGCAT CTCTTTGAAT GGCGCTGGGC CGACATTGCA CTGGAGTGC G AACACTATTT	1800
AGCTCCTAAT GGGTTGGAG GAGTCAGGT TTCTCCTCCA AATGAAAACA TTGTCATTAC	1860
TAATCCGAAC AGGCCCTGGT GGGAAAGATA CCAGCCCATC AGCTACAAGA TCTGCAGTCG	1920
ATCGGGCAAT GAAAATGAAT TCAGAGACAT GGTGACCAGA TGCAACAATG TTGGAGTTCG	1980
TATTTATGTG GATGCTGTTG TCAATCACAT GTGTGGATCT ATGGGTGGCA CGGGCACCCA	2040

CTCAACATGT GGGAGCTATT TCAACACCGG GACTAGAGAT TTTCCCGCTG TGCGTACTC	2100
TGCCTGGGAT TTCAATGACG GCAAATGTCA CACTGCAAGT GGAGACATCG AAAATTATGG	2160
GGACATGTAT CAGGTCCGGG ATTGCAAGTT GTCCAGCCTT CTTGATCTGG CTCTGGAGAA	2220
GGACTATGTA CGCTCAACAA TTGCAGCGTA CATGAATCAC CTCATTGATA TGGGTGTAGC	2280
AGGGTTCCGG ATCGATGCTG CCAAGCATAT GTGGCCAGGG GACATAAGAG CGTTTCTGGA	2340
CAAACGTGAC GATCTAAATA CTCAGTGGTT TTCAGCAGGA ACGAAACCCCT TTATTTACCA	2400
AGAGGTAATT GACTTGGGAG GAGAGCCAAT CACAGGCAGT CAGTACTTTG GGAATGGCCG	2460
CGTGACAGAA TTCAAGTATG GTGCCAAACT GGGGACGGTG ATCCGGAAGT GGAATGGAGA	2520
GAAGATGGCC TACTTAAAGA ACTGGGGAGA AGGCTGGGC TTTGTGCCCTT CTGACAGAGC	2580
CCTGGTGTGTT GTGGATAACC ACGACAACCA GCGGGGGCAC GGGGCAGGCG GAGCTTCCAT	2640
TCTTACTTTTC TGGGATGCCA GGCTTTATAA AATGGCGGTT GGTTTCATGC TCGCTCATCC	2700
GTACGGGTTTC ACACGGGTGA TGTCAAGTTA TCGTTGGCCA AGATATTCG AAAACGGAGT	2760
GGATGTTAAC GACTGGGTGG GACCACCAAG TAACTCGGAC GGATCGACGA AGTCCGTTAC	2820
AATCAACGCA GACACTACCT GTGGCAATGA CTGGGTCTGC GAACATCGCT GGGCACAAAT	2880
AAGGAACATG GTTATCTTCC GTAATGTGGT AGACGGTCAG CCTTTCTCAA ACTGGTGGGA	2940
CAACGGGAGC AATCAAGTAG CTTCGGTGCG CGGCGACAGA GGCTTCATTG TCTTTAATAA	3000
TGATGACTGG TATATGAATG TCGATTGCA AACTGGTCTG CCTGCTGGAA CCTACTGCGA	3060
TGTTATTTCT GGACAAAAGG AAGGCAGTGC GTGTACTGGA AAGCAGGTGT ACGTTTCCTC	3120
GGATGGAAAG GCCAATTCC AGATTAGTAA CAGCGATGAA GATCCATTG TTGCAATTCA	3180
CGTTGATGCC AAGTTATAAG CTTCGAGGAT CCACTAGTAA CGGCCGCCAG TGTGCTGGAA	3240
TCGGCTTGT CGACATCTAG GGCGCCAAT TCCGCCCTC TCCCCCCCCC CCCTAACGTT	3300
ACTGGCCGAA GCCGCTTGGAA ATAAGGCCGG TGTGTGTTTG TCTATATGTG ATTTTCCACC	3360
ATATTGCCGT CTTTGGCAA TGTGAGGCC CGGAAACCTG GCCCTGTCTT CTTGACGAGC	3420
ATTCCTAGGG GTCTTCCCC TCTGCCAAA GGAATGCAAG GTCTGTTGAA TGTGCGTGAAG	3480
GAAGCAGTTC CTCTGGAAGC TTCTTGAAGA CAAACAAACGT CTGTAGCGAC CCTTTGCAGG	3540
CAGCGGAACC CCCCACCTGG CGACAGGTGC CTCTGCGGCC AAAAGCCACG TGTATAAGAT	3600
ACACCTGCAA AGGCAGGCACA ACCCCAGTGC CACGTTGTGA GTTGGATAAGT TGTGGAAAGA	3660
GTCAAATGGC TCTCCTCAAG CGTAGTCAAC AAGGGGCTGA AGGATGCCA GAAGGTACCC	3720

CATTGTATGG	GAATCTGATC	TGGGGCCTCG	GTGCACATGC	TTTACATGTG	TTTAGTCGAG	3780
GTTAAAAAAAG	CTCTAGGCC	CCCGAACAC	GGGGACGTGG	TTTCCTTTG	AAAAACACGA	3840
TGATAAGCTT	GCCACAACCC	AAACAGCGTC	AACAGCGTGC	CGCAGATCCC	GGGCAATGAG	3900
ATATGAAAAA	GCCTGAACTC	ACCGCGACGT	CTGTCGAGAA	GTTTCTGATC	GAAAAGTTCG	3960
ACAGCGTCTC	CGACCTGATG	CAGCTCTCGG	AGGGCGAAGA	ATCTCGTGC	TTCAAGCTTCG	4020
ATGTTAGGAGG	GCGTGGATAT	GTCCTGCGGG	TAAATAGCTG	CGCCGATGGT	TTCTACAAAG	4080
ATCGTTATGT	TTATCGGCAC	TTTGCATCGG	CCGCGCTCCC	GATTCCGGAA	GTGCTTGACA	4140
TTGGGAAATT	CAGCGAGAGC	CTGACCTATT	GCATCTCCCG	CCGTGCACAG	GGTGTACACGT	4200
TGCAAGACCT	GCCTGAAACC	GAACTGCCCG	CTGTTCTGCA	GCCGGTCGCG	GAGGCCATGG	4260
ATGCGATCGC	TGCGGCCGAT	CTTAGCCAGA	CGAGCGGGTT	CGGCCCATTC	GGACCGCAAG	4320
GAATCGGTCA	ATACACTACA	TGGCGTGATT	TCATATGCGC	GATTGCTGAT	CCCCATGTGT	4380
ATCACTGGCA	AACTGTGATG	GACGACACCG	TCAGTGCCTC	CGTCGCGCAG	GCTCTCGATG	4440
AGCTGATGCT	TTGGGCCGAG	GACTGCCCG	AACTCCGGCA	CCTCGTGCAC	GCGGATTTCG	4500
GCTCCAACAA	TGTCCTGACG	GACAATGGCC	GCATAACAGC	GGTCATTGAC	TGGAGCGAGG	4560
CGATGTTCGG	GGATTCCCAA	TACCGAGGTG	CCAACATCTT	CTTCTGGAGG	CCGTGGTTGG	4620
CTTGTATGGA	GCAGCAGACG	CGCTACTTCG	AGCGGAGGCA	TCCGGAGCTT	GCAGGATCGC	4680
CGCGGCTCCG	GGCGTATATG	CTCCGCATTG	GTCTTGACCA	ACTCTATCAG	AGCTTGGTTG	4740
ACGGCAATT	CGATGATGCA	GCTTGGCGC	AGGGTCGATG	CGACGCAATC	GTCCGATCCG	4800
GAGCCGGGAC	TGTCGGCGT	ACACAAATCG	CCCGCAGAAG	CGCGGCCGTC	TGGACCGATG	4860
GCTGTGTAGA	AGTACTCGCC	GATAGTGGAA	ACCGACGCC	CAGCACTCGT	CCGAGGGCAA	4920
AGGAATAGGG	GAGATGGGGG	AGGCTAACTG	AAACACGGAA	GGGCCCCGG	GACTCTGGGG	4980
TTCGATAAAA	TAAAAGATT	TATTTAGTCT	CCAGAAAAAG	GGGGGAATGA	AAGACCCCAC	5040
CTGTAGGTTT	GGCAAGCTAG	CTTAAGTAAC	GCCATTTGC	AAGGCATGGA	AAAATACATA	5100
ACTGAGAATA	GAGAAGTTCA	GATCAAGGTC	AGGAACAGAT	GGAACAGCTG	AATATGGGCC	5160
AAACAGGATA	TCTGTGGTAA	GCAGTTCTG	CCCCGGCTCA	GGGCCAAGAA	CAGATGGAAC	5220
AGCTGAATAT	GGGCCAAACA	GGATATCTGT	GGTAAGCAGT	TCCTGCCCG	GCTCAGGGCC	5280
AAGAACAGAT	GGTCCCCAGA	TGCGGTCCAG	CCCTCAGCAG	TTTCTAGAGA	ACCATCAGAT	5340
GTTCAGGG	TGCCCAAGG	ACCTGAAATG	ACCCCTGTGCC	TTATTTGAAC	TAACCAATCA	5400

GTTCGCTTCT CGCTTCTGTT CGCGCGCTTC TGCTCCCCGA GCTCAATAAA AGAGCCCACA	5460
ACCCCTCACT CGGGGCGCCA GTCCCTCCGAT TGACTGAGTC GCCCGGGTAC CCGTGTATCC	5520
AATAAACCTT CTTGCAGTTG CATCCGACTT GTGGTCTCGC TGTTCCCTGG GAGGGTCTCC	5580
TCTGAGTGAT TGACTACCCG TCAGCGGGGG TCTTTCATTT GGGGGCTCGT CCGGGATCGG	5640
GAGACCCCTG CCCAGGGACC ACCGACCCAC CACCGGGAGG TAAGCTGGCT GCCTCGCG	5700
TTTCGGTGAT GACGGTGAAA ACCTCTGACA CATGCAGCTC CCGGAGACGG TCACAGCTTG	5760
TCTGTAAGCG GATGCCGGGA GCAGACAAGC CCGTCAGGGC GCGTCAGCGG GTGTTGGCGG	5820
GTGTCGGGCG GCAGCCATGA CCCAGTCACG TAGCGATAGC GGAGTGTATA CTGGCTTAAC	5880
TATGCGGCAT CAGAGCAGAT TGTACTGAGA GTGCACCATA TGCGGTGTGA AATACCGCAC	5940
AGATGCGTAA GGAGAAAATA CCGCATCAGG CGCTCTTCCG CTTCCCTCGCT CACTGACTCG	6000
CTGCGCTCGG TCGTTGGCT GCGCGAGCG GTATCAGCTC ACTCAAAGGC GGTAATACGG	6060
TTATCCACAG AATCAGGGGA TAACGCAGGA AAGAACATGT GAGCAAAAGG CCAGCAAAAG	6120
GCCAGGAACC GTAAAAAGGC CGCGTTGCTG GCGTTTTCC ATAGGCTCCG CCCCCCTGAC	6180
GAGCATCACA AAAATCGACG CTCAAGTCAG AGGTGGCGAA ACCCGACAGG ACTATAAAGA	6240
TACCAGGCCT TTCCCCCTGG AAGCTCCCTC GTGCGCTCTC CTGTTCCGAC CCTGCCGCTT	6300
ACCGGATACC TGTCCGCCCT TCTCCCTCG GGAAGCGTGG CGCTTCTCA TAGCTCACGC	6360
TGTAGGTATC TCAGTTGGT GTAGGTCGTT CGCTCCAAGC TGGGCTGTGT GCACGAACCC	6420
CCCGTTTCAGC CCGACCGCTG CGCCTTATCC GGTAACTATC GTCTTGAGTC CAACCCGGTA	6480
AGACACGACT TATGCCACT GGCAGCAGCC ACTGGTAACA GGATTAGCAG AGCGAGGTAT	6540
GTAGGCGGTG CTACAGAGTT CTTGAAGTGG TGGCCTAACT ACGGCTACAC TAGAAGGACA	6600
GTATTTGGTA TCTGCGCTCT GCTGAAGCCA GTTACCTTCG GAAAAAGAGT TGGTAGCTCT	6660
TGATCCGGCA AACAAACCAC CGCTGGTAGC GGTGGTTTT TTGTTGCAA GCAGCAGATT	6720
ACGCGCAGAA AAAAAGGATC TCAAGAAGAT CCTTGATCT TTTCTACGGG GTCTGACGCT	6780
CAGTGGAACG AAAACTCACG TTAAGGGATT TTGGTCATGA GATTATCAAA AAGGATCTTC	6840
ACCTAGATCC TTTAAATTA AAAATGAAGT TTTAAATCAA TCTAAAGTAT ATATGAGTAA	6900
ACTTGGTCTG ACAGTTACCA ATGCTTAATC AGTGAGGCAC CTATCTCAGC GATCTGTCTA	6960
TTTCGTTCAT CCATAGTTGC CTGACTCCCC GTCGTGTAGA TAACTACGAT ACGGGAGGGC	7020
TTACCATCTG GCCCCAGTGC TGCAATGATA CCGCGAGACC CACGCTCACC GGCTCCAGAT	7080

TTATCAGCAA TAAACCAGCC AGCCGGAAGG GCCGAGCGCA GAAAGTGGTCC TGCAACTTTA	7140
TCCGCCTCCA TCCAGTCTAT TAATTGTTGC CGGGAAGCTA GAGTAAGTAG TTCGCCAGTT	7200
AATAGTTGC GCAACGTTGT TGCCATTGCT GCAGGCATCG TGGTGTACCG CTCGTCGTTT	7260
GGTATGGCTT CATTCAAGCTC CGGTTCCCAA CGATCAAGGC GAGTTACATG ATCCCCCATG	7320
TTGTGCAAAA AAGCGGTTAG CTCCCTTCGGT CCTCCGATCG TTGTCAGAAC TAAGTTGGCC	7380
GCAGTGTTAT CACTCATGGT TATGGCAGCA CTGCATAATT CTCTTACTGT CATGCCATCC	7440
GTAAGATGCT TTTCTGTGAC TGGTGAGTAC TCAACCAAGT CATTCTGAGA ATAGTGTATG	7500
CGGCGACCGA GTTGCTCTTG CCCGGCGTCA ACACGGGATA ATACCGCGCC ACATAGCAGA	7560
ACTTTAAAAG TGCTCATCAT TGAAAACGT TCTTCGGGGC GAAAACCTCTC AAGGATCTTA	7620
CCGCTGTTGA GATCCAGTTC GATGTAACCC ACTCGTGCAC CCAACTGATC TTCAGCATCT	7680
TTTACTTTCA CCAGCGTTTC TGGGTGAGCA AAAACAGGAA GGCAAAATGC CGAAAAAAAG	7740
GGAATAAGGG CGACACGGAA ATGTTGAATA CTCATACTCT TCCTTTTCA ATATTATTGA	7800
AGCATTATTC AGGGTTATTG TCTCATGAGC GGATACATAT TTGAATGTAT TTAGAAAAAT	7860
AAACAAATAG GGGTTCCGCG CACATTCCC CGAAAAGTGC CACCTGACGT CTAAGAAACC	7920
ATTATTATCA TGACATTAAC CTATAAAAAT AGGCGTATCA CGAGGCCCTT TCGTCTTCAA	7980

## (2) INFORMATION FOR SEQ ID NO: 3:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 7311 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: circular

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

GAATTCAAC CAGATCACCG AAAACTGTCC TCCAAATGTG TCCCCCTCAC ACTCCCAAAT	60
TCGCGGGCTT CTGCCTCTTA GACCACTCTA CCCTATTCCC CACACTCACC GGAGCCAAAG	120
CCGCGGGCCCT TCCGTTCTT TGCTTTGAA AGACCCCACC CGTAGGTGGC AAGCTAGCTT	180
AAGTAACGCC ACTTTGCAAG GCATGGAAAA ATACATAACT GAGAATAGAA AAGTTCAGAT	240
CAAGGTCAGG AACAAAGAAA CAGCTGAATA CCAACAGGA TATCTGTGGT AAGCGGTTCC	300

TGCCCGGGCT CAGGGCCAAG AACAGATGAG ACAGCTGAGT GATGGGCCAA ACAGGATATC	360
TGTGGTAAGC AGTTCTGCC CCGGCTCGGG GCCAAGAACAA GATGGTCCCC AGATGCGGTC	420
CAGCCCTCAG CAGTTCTAG TGAATCATCA GATGTTCCA GGGTGCCCCA AGGACCTGAA	480
AATGACCCCTG TACCTTATTT GAACTAACCA ATCAGTTCGC TTCTCGCTTC TGTCGCGCG	540
CTTCCGCTCT CCGAGCTCAA TAAAAGAGCC CACAACCCCT CACTCGGCGC GCCAGTCTTC	600
CGATAGACTG CGTCGCCCCG GTACCCGTAT TCCCAATAAA GCCTCTTGCT GTTTGCATCC	660
GAATCGTGGT CTCGCTGTT CTTGGGAGGG TCTCCCTCTGA GTGATTGACT ACCCACGACG	720
GGGGTCTTTC ATTTGGGGGC TCGTCCGGGA TTTGGAGACC CCTGCCAGG GACCACCGAC	780
CCACCACCGG GAGGTAAGCT GGCCAGCAAC TTATCTGTGT CTGTCCGATT GTCTAGTGTGTC	840
TATGTTTGAT GTTATGCGCC TCGTCTGTA CTAGTTAGCT AACTAGCTCT GTATCTGGCG	900
GACCCGTGGT GGAACTGACG AGTTCTGAAC ACCCGGCCGC AACCCCTGGGA GACGTCCCAG	960
GGACTTTGGG GGCGTTTTT GTGGCCCGAC CTGAGGAAGG GAGTCGATGT GGAATCCGAC	1020
CCCGTCAGGA TATGTGGTTC TGGTAGGAGA CGAGAACCTA AAACAGTTCC CGCCTCCGTC	1080
TGAATTTTG CTTTCGGTTT GGAACCGAAG CCGCGCGTCT TGTCTGCTGC AGCGCTGCAG	1140
CATCGTTCTG TGGTGTCTCT GTCTGACTGT GTTTCTGTAT TTGTCTGAAA ATTAGGGCCA	1200
GAATGTTACC ACTCCCTAA GTTGACCTT AGGTCACTGG AAAGATGTCG AGCGGATCGC	1260
TCACAACCAG TCGGTAGATG TCAAGAACAGAG ACAGTTGGGTT ACCTTCTGCT CTGCAGAACATG	1320
GCCAACCTTT AACGTCGGAT GGCGCGAGA CGGCACCTTT AACCGAGACC TCATCACCCA	1380
GGTTAAGATC AAGGTCTTTT CACCTGGCCC GCATGGACAC CCAGACCAGG TCCCCTACAT	1440
CGTGACCTGG GAAGCCTTGG CTTTGACCC CCCTCCCTGG GTCAAGCCCT TTGTACACCC	1500
TAAGCCTCCG CCTCCTCTTC CTCCATCCGC CCCGTCTCTC CCCCTTGAAC CTCCTCGTTC	1560
GACCCCGCCT CGATCCTCCC TTTATCCAGC CCTCACTCCT TCTCTAGGCG CGGAAATTCC	1620
GATCTGATCA AGAGACAGGA TGAGGATCGT TTCGCATGAT TGAACAAAGAT GGATTGCACG	1680
CAGGTTCTCC GGCGCGTTGG GTGGAGAGGC TATTGGCTA TGACTGGCA CAACAGACAA	1740
TCGGCTGCTC TGATGCCGCC GTGTTCCGGC TGTCAGCGCA GGGGCGCCCG GTTCTTTTG	1800
TCAAGACCGA CCTGTCCGGT GCCCTGAATG AACTGCAGGA CGAGGCAGCG CGGCTATCGT	1860
GGCTGGCCAC GACGGGCAGTT CCTTGCGCAG CTGTGCTCGA CGTTGTCACT GAAGCGGGAA	1920
GGGACTGGCT GCTATTGGGC GAAGTGCCGG GGCAGGATCT CCTGTCATCT CACCTTGCTC	1980

CTGCCGAGAA AGTATCCATC ATGGCTGATG CAATGCCCG GCTGCATACG CTTGATCCGG	2040
CTACCTGCC ATTGACCAAC CAAGCGAAC ATCGCATCGA GCGAGCACGT ACTCGGATGG	2100
AAGCCGGTCT TGTGATCTAG GATGATCTGG ACGAAGAGCA TCAGGGCTC GCGCCAGCCG	2160
AACTGTTCGC CAGGCTCAAG GCGCGATGC CCGACGGCGA GGATCTCGTC GTGACCCATG	2220
GCGATGCCCTG CTTGCCAAT ATCATGGTGG AAAATGGCCG CTTTTCTGGA TTCATCGACT	2280
GTGGCCGGCT GGGTGTGGCG GACCGCTATC AGGACATAGC GTTGGCTACC CGTGATATTG	2340
CTGAAGAGCT TGGCGCGAA TGGGCTGACC GCTTCCTCGT GCTTTACGGT ATCGCCGCTC	2400
CCGATTGCA GCGCATCGCC TTCTATCGCC TTCTTGACGA GTTCTTCTGA GCGGGACTCT	2460
GGGGTTCGAA ATGACCGACC AAGCGACGCC CAACCTGCCA TCACGAGATT TCGATTCCAC	2520
CGCCGCCTTC TATGAAAGGT TGGGCTTCGG AATCGTTTTC CGGGACGCCG GCTGGATGAT	2580
CCTCCAGCGC GGGGATCTCA TGCTGGAGTT CTTCGCCAC CCCGGGCTCG ATCCCCTCGC	2640
GAGTTGGTTC AGCTGCTGCC TGAGGCTGGA CGACCTCGCG GAGTTCTACC GGCAGTGCAA	2700
ATCCGTCGGC ATCCAGGAAA CCAGCAGCGG CTATCCGCGC ATCCATGCC CCGAACTGCA	2760
GGAGTGGGGA GGCACGATGG CCGCTTTGGT CGAGGCGGAT CCGGGCAGAA ATGGTTGAAC	2820
TCCCCGAGAGT GTCCTACACC TAGGGGAGAA GCAGCCAAGG GGTTGTTTCC CACCAAGGAC	2880
GACCCGTCTG CGCACAAACG GATGAGCCA TCAGACAAAG ACATATTCA TCTCTGCTGC	2940
AAACTTGGCA TAGCTCTGCT TTGCTGGGG CTATTGGGGG AAGTTGCGGT TCGTGCTCGC	3000
AGGGCTCTCA CCCTTGACTC TTTTAATAGC TCTTCTGTGC AAGATTACAA TCTAAACAAT	3060
TCGGAGAACT CGACCTTCCT CCTGAGGCAA GGACCACAGC CAACTTCCTC TTACAAGCCG	3120
CATCGATTTT GTCCTTCAGA AATAGAAATA AGAATGCTTG CTAAAAATTA TATTTTACCC	3180
AATAAGACCA ATCCAATAGG TAGATTATTA GTTACTATGT TAAGAAATGA ATCATTATCT	3240
TTTAGTACTA TTTTTACTCA AATTCAAGAG TTAGAAATGG GAATAGAAAA TAGAAAGAGA	3300
CGCTCAACCT CAATTGAAGA ACAGGTGCAA GGACTATTGA CCACAGGCCT AGAAGTAAAA	3360
AAGGGAAAAAA AGAGTGTGTT TGTCAAAATA GGAGACAGGT GGTGGCAACC AGGGACTTAT	3420
AGGGGACCTT ACATCTACAG ACCAACAGAT GCCCCCTTAC CATATACAGG AAGATATGAC	3480
TTAAATTGGG ATAGGTGGGT TACAGTCAAT GGCTATAAAG TGTTATATAG ATCCCTCCCT	3540
TTTCGTGAAA GACTCGCCAG AGCTAGACCT CCTTGGTGTGTA TGTTGTCTCA AGAAGAAAAAA	3600
GACGACATGA AACAAACAGGT ACATGATTAT ATTTATCTAG GAACAGGAAT GCACTTTGG	3660

GGAAAGATT TCCATACCAA GGAGGGACA GTGGCTGGAC TAATAGAAC	TTATTCTGCA	3720
AAAACTCATG GCATGAGTTA TTATGAATAG CCTTTATTGG CCCAACCTTG CGGTTCCCAG		3780
GGCTTAAGTA AGTTTTGGT TACAAACTGT TCTTAAACG AGGATGTGAG ACAAGTGGTT		3840
TCCTGACTTG GTTTGGTATC AAAGGTTCTG ATCTGAGCTC TGAGTGTCT ATTTCCTAT		3900
GTTCTTTGG AATTATCCA AATCTTATGT AAATGCTTAT GTAAACCAAG ATATAAAAGA		3960
GTGCTGATTT TTTGAGTAAA CTTGCAACAG TCCTAACATT CACCTCTTGT GTGTTGTGT		4020
CTGTTGCCA TCCCCTCTCC GCTCGTCACT TATCCTTCAC TTTCCAGAGG GTCCCCCGC		4080
AGACCCCGGC GACCCTCAGG TCGGCCGACT GCGGCAGCTG GCGCCCGAAC AGGGACCCTC		4140
GGATAAGTGA CCCTTGTCTC TATTCTACT ATTTGGTGTGTT TGTCTTGTAT TGTCTCTTTC		4200
TTGTCTGGCT ATCATCACAA GAGCGGAACG GACTCACCAT AGGGACCAAG CTTGTCGACA		4260
TTTCTGCAGA TATCCATCAC ACTGGCGGCC GCTCGAGCAT GCATCTAGAA CATCGATAAA		4320
ATAAAAAGATT TTATTTAGTC TCCAGAAAAA GGGGGGAATG AAAGACCCCA CCTGTAGGTT		4380
TGGCAAGCTA GCTTAAGTAA CGCCATTTTG CAAGGCATGG AAAAATACAT AACTGAGAAT		4440
AGAGAAGTTC AGATCAAGGT CAGGAACAGA TGGAACAGCT GAATATGGC CAAACAGGAT		4500
ATCTGTGGTA AGCAGTTCCCT GCCCCGGCTC AGGGCCAAGA ACAGATGGAA CAGCTGAATA		4560
TGGGCCAAAC AGGATATCTG TGGTAAGCAG TTCCTGCCCG GGCTCAGGGC CAAGAACAGA		4620
TGGTCCCCAG ATGCGGTCCA GCCCTCAGCA GTTTCTAGAG AACCATCAGA TGTTCCAGG		4680
GTGCCCAAG GACCTGAAAT GACCTGTGC CTTATTTGAA CTAACCAATC AGTCGCTTC		4740
TCGCTTCTGT TCGCGCGCTT CTGCTCCCCG AGCTCAATAA AAGAGCCAC AACCCCTCAC		4800
TCGGGGCGCC AGTCCTCCGA TTGACTGAGT CGCCCGGGTA CCCGTGTATC CAATAAACCC		4860
TCTTGCAGTT GCATCCGACT TGTGGTCTCG CTGTTCTTG GGAGGGTCTC CTCTGAGTGA		4920
TTGACTACCC GTCAGCGGGG GTCTTTCATT TGGGGGCTCG TCCGGGATCG GGAGACCCCT		4980
GCCCAGGGAC CACCGACCCA CCACCGGGAG GTAAGCTGGC TGCTCGCGC GTTTCGGTGA		5040
TGACGGTGAA AACCTCTGAC ACATGCAGCT CCCGGAGACG GTCACAGCTT GTCTGTAAGC		5100
GGATGCCGGG AGCAGACAAG CCCGTCAGGG CGCGTCAGCG GGTGTTGGCG GGTGTCGGGG		5160
CGCAGCCATG ACCCAGTCAC GTAGCGATAG CGGAGTGTAT ACTGGCTTAA CTATGCGCA		5220
TCAGAGCAGA TTGTACTGAG AGTGCACCAT ATGCGGTGTG AAATACCGCA CAGATGCGTA		5280
AGGAGAAAAT ACCGCATCAG GCGCTCTTCC GCTTCCTCGC TCACTGACTC GCTGCGCTCG		5340

GTCGTTCGGC	TGCGGGCGAGC	GGTATCAGCT	CACTCAAAGG	CGGTAATAACG	GTTATCCACA	5400
GAATCAGGGG	ATAACGCAGG	AAAGAACATG	TGAGCAAAAG	GCCAGCAAA	GGCCAGGAAC	5460
CGTAAAAAGG	CCGCGTTGCT	GGCGTTTTTC	CATAGGCTCC	GCCCCCTGA	CGAGCATCAC	5520
AAAAATCGAC	GCTCAAGTCA	GAGGTGGCGA	AACCCGACAG	GACTATAAAG	ATACCAGGCG	5580
TTTCCCCCTG	GAAGCTCCCT	CGTGCCTCT	CCTGTTCCGA	CCCTGCCGCT	TACCGGATAC	5640
CTGTCCGCCT	TTCTCCCTTC	GGGAAGCGTG	GCGCTTTCTC	ATAGCTCACG	CTGTAGGTAT	5700
CTCAGTTCGG	TGTAGGTCGT	TCGCTCCAAG	CTGGGCTGTG	TGCACGAACC	CCCCGTTTAG	5760
CCCCGACCGCT	GCGCCTTATC	CGGTAACTAT	CGTCTTGAGT	CCAACCCGGT	AAGACACGAC	5820
TTATGCCAC	TGGCAGCAGC	CACTGGTAAC	AGGATTAGCA	GAGCGAGGTA	TGTAGGCGGT	5880
GCTACAGAGT	TCTTGAAGTG	GTGGCCTAAC	TACGGCTACA	CTAGAAGGAC	AGTATTTGGT	5940
ATCTGCCTC	TGCTGAAGCC	AGTTACCTTC	GGAAAAAGAG	TTGGTAGCTC	TTGATCCGGC	6000
AAACAAACCA	CCGCTGGTAG	CGGTGGTTTT	TTTGGTTGCA	AGCAGCAGAT	TACGCGCAGA	6060
AAAAAAGGAT	CTCAAGAAGA	TCCTTGATC	TTTCTACGG	GGTCTGACGC	TCAGTGGAAC	6120
GAAAACCTCAC	GTTAAGGGAT	TTTGGTCATG	AGATTATCAA	AAAGGATCTT	CACCTAGATC	6180
CTTTTAAATT	AAAAATGAAG	TTTAAATCA	ATCTAAAGTA	TATATGAGTA	AACTGGTCT	6240
GACAGTTACC	AATGCTTAAT	CAGTGAGGCA	CCTATCTCAG	CGATCTGTCT	ATTCGTTCA	6300
TCCATAGTTG	CCTGACTCCC	CGTCGTGTAG	ATAACTACGA	TACGGGAGGG	CTTACCATCT	6360
GGCCCCAGTG	CTGCAATGAT	ACCGCGAGAC	CCACGCTCAC	CGGCTCCAGA	TTTATCAGCA	6420
ATAAACCCAGC	CAGCCGGAAG	GGCCGAGCGC	AGAAGTGGTC	CTGCAACTTT	ATCCGCCTCC	6480
ATCCAGTCTA	TTAATTGTTG	CCGGGAAGCT	AGAGTAAGTA	GTTCGCCAGT	TAATAGTTG	6540
CGCAACGTTG	TTGCCATTGTC	TGCAGGCATC	GTGGTGTAC	GCTCGTCGT	TGGTATGGCT	6600
TCATTCAAGCT	CCGGTTCCCA	ACGATCAAGG	CGAGTTACAT	GATCCCCAT	GTTGTGCAAA	6660
AAAGCGGTTA	GCTCCTTCGG	TCCTCCGATC	GTTGTCAGAA	GTAAGTTGGC	CGCAGTGTAA	6720
TCACTCATGG	TTATGGCAGC	ACTGCATAAT	TCTCTTACTG	TCATGCCATC	CGTAAGATGC	6780
TTTTCTGTGA	CTGGTGAGTA	CTCAACCAAG	TCATTCTGAG	AATAGTGTAT	CGGGCGACCG	6840
AGTTGCTCTT	GCCCCGGCGTC	AACACGGGAT	AATACCGCGC	CACATAGCAG	AACTTTAAAA	6900
G TGCTCATCA	TTGGAAAACG	TTCTTCGGGG	CGAAAACCTCT	CAAGGATCTT	ACCGCTGTTG	6960
AGATCCAGTT	CGATGTAACC	CACTCGTGCA	CCCAACTGAT	CTTCAGCATC	TTTACTTTC	7020

40

ACCAGCGTTT CTGGGTGAGC AAAAACAGGA AGGCAAAATG CCGCAAAAAA GGGAAATAAGG	7080
GCGACACGGA AATGTTGAAT ACTCATACTC TTCCCTTTTC AATATTATTG AAGCATTTAT	7140
CAGGGTTATT GTCTCATGAG CGGATACATA TTTGAATGTA TTTAGAAAAA TAAACAAATA	7200
GGGGTTCCGC GCACATTTC CCGAAAAGTG CCACCTGACG TCTAAGAAC CATTATTATC	7260
ATGACATTAAC CCTATAAAAAA TAGGCGTATC ACGAGGCCCT TTCGTCTTCA A	7311

## (2) INFORMATION FOR SEQ ID NO: 4:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 7885 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: circular

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

GAATTCTACAC CAGATCACCG AAAACTGTCC TCCAAATGTG TCCCCCTCAC ACTCCCAAAT	60
TCGCAGGGCTT CTGCCTCTTA GACCACTCTA CCCTATTCCC CACACTCACC GGAGCCAAAG	120
CCGCAGGGCCCT TCCGTTCTT TGCTTTGAA AGACCCCACC CGTAGGTGGC AAGCTAGCTT	180
AAGTAACGCC ACTTTGCAAG GCATGGAAAA ATACATAACT GAGAATAGAA AAGTTCAGAT	240
CAAGGTCAGG AACAAAGAAA CAGCTGAATA CCAAACAGGA TATCTGTGGT AAGCGGTTCC	300
TGCCCTGGCT CAGGGCCAAG AACAGATGAG ACAGCTGAGT GATGGGCCAA ACAGGATATC	360
TGTGGTAAGC AGTTCTGCC CCGGCTCGGG GCCAAGAACAA GATGGTCCCC AGATGCGGTC	420
CAGCCCTCAG CAGTTCTAG TGAATCATCA GATGTTCCA GGGTGCCCCA AGGACCTGAA	480
AATGACCTTG TACCTTATTT GAACTAACCA ATCAGTTCGC TTCTCGCTTC TGTTCGCGCG	540
CTTCCGCTCT CCGAGCTCAA TAAAAGAGCC CACAACCCCT CACTCGGCGC GCCAGTCTTC	600
CGATAGACTG CGTCGCCGG GTACCCGTAT TCCCAATAAA GCCTCTTGCT GTTTGCATCC	660
GAATCGTGGT CTCGCTGTT CTTGGGAGGG TCTCCTCTGA GTGATTGACT ACCCACGACG	720
GGGGCTTTC ATTTGGGGC TCGTCCGGGA TTTGGAGACC CCTGCCAGG GACCACCGAC	780
CCACCACCGG GAGGTAAGCT GGCCAGCAAC TTATCTGTGT CTGTCCGATT GTCTAGTGT	840
TATGTTTGAT GTTATGCGCC TGCCTCTGTA CTAGTTAGCT AACTAGCTCT GTATCTGGCG	900
GACCCGTGGT GGAACGTGACG AGTTCTGAAC ACCCGGCCGC AACCCCTGGGA GACGTCCCAG	960

GGACTTTGGG GGCGTTTT GTGGCCGAC CTGAGGAAGG GAGTCGATGT GGAATCCGAC	1020
CCCGTCAGGA TATGTGGTTC TGGTAGGAGA CGAGAACCTA AAACAGTTCC CGCCTCCGTC	1080
TGAATTGGT CTTTCGGTTT GGAACCGAAG CCGCGCGTCT TGTCTGCTGC AGCGCTGCAG	1140
CATCGTTCTG TGTTGCTCT GTCTGACTGT GTTTCTGTAT TTGTCTGAAA ATTAGGGCCA	1200
GACTGTTACC ACTCCCTAA GTTGACCTT AGGTCACTGG AAAGATGTCG AGCGGATCGC	1260
TCACAACCAAG TCGGTAGATG TCAAGAAGAG ACGTTGGGTT ACCTTCTGCT CTGCAGAATG	1320
GCCAACCTTT AACGTCGGAT GGCGCGAGA CGGCACCTT AACCGAGACC TCATCACCCA	1380
GGTTAAGATC AAGGTCTTT CACCTGGCCC GCATGGACAC CCAGACCAGG TCCCCTACAT	1440
CGTGACCTGG GAAGCCTTGG CTTTGACCC CCCTCCCTGG GTCAAGCCCT TTGTACACCC	1500
TAAGCCTCCG CCTCCTCTTC CTCCATCCGC CCCGTCTCTC CCCCTTGAAC CTCTCGTTC	1560
GACCCCGCCT CGATCCTCCC TTTATCCAGC CCTCACTCCT TCTCTAGGCG CCGGAATTCC	1620
GATCTGATCA AGAGACAGGA TGAGGATCGT TTGCGATGAT TGAACAAGAT GGATTGCACG	1680
CAGGTTCTCC GGCGCTTGG GTGGAGAGGC TATTCGGCTA TGACTGGCA CAACAGACAA	1740
TCGGCTGCTC TGATGCCGCC GTGTTCCGGC TGTCAGCGCA GGGGCGCCCG GTTCTTTTG	1800
TCAAGACCGA CCTGTCCGGT GCCCTGAATG AACTGCAGGA CGAGGCAGCG CGGCTATCGT	1860
GGCTGGCCAC GACGGCGTT CCTTGCGCAG CTGTGCTCGA CGTTGTCACT GAAGCGGGAA	1920
GGGACTGGCT GCTATTGGGC GAAATGCCGG GGCAGGATCT CCTGTCATCT CACCTTGCTC	1980
CTGCCGAGAA AGTATCCATC ATGGCTGATG CAATGCCGG GCTGCATACG CTTGATCCGG	2040
CTACCTGCC ATTGACCCAC CAAGCGAAAC ATCGCATCGA GCGAGCACGT ACTCGGATGG	2100
AAGCCGGTCT TGTCGATCAG GATGATCTGG ACGAAGAGCA TCAGGGCTC GCGCCAGCCG	2160
AACTGTTCGC CAGGCTCAAG GCGCGCATGC CCGACGGCGA GGATCTCGTC GTGACCCATG	2220
GCGATGCCCTG CTTGCCGAAT ATCATGGTGG AAAATGGCCG CTTTCTGGA TTCATCGACT	2280
GTGGCCGGCT GGGTGTGGCG GACCGCTATC AGGACATAGC GTTGGCTACC CGTGATATTG	2340
CTGAAGAGCT TGGCGGGCAA TGGGCTGACC GCTTCCTCGT GCTTTACGGT ATCGCCGCTC	2400
CCGATTGCA GCGCATCGCC TTCTATCGCC TTCTTGACGA GTTCTTCTGA GCGGGACTCT	2460
GGGGTTCGAA ATGACCGACC AAGCGACGCC CAACCTGCCA TCACGAGATT TCGATTCCAC	2520
CGCCGCCCTTC TATGAAAGGT TGGGCTTCGG AATCGTTTTC CGGGACGCCG GCTGGATGAT	2580
CCTCCAGCGC GGGGATCTCA TGCTGGAGTT CTTCGCCCCAC CCCGGGCTCG ATCCCCCTCGC	2640

GAGTTGGTTC AGCTGCTGCC TGAGGCTGGA CGACCTCGCG GAGTTCTACC GGCAGTGCAA	2700
ATCCGTCGGC ATCCAGGAAA CCAGCAGCGG CTATCCGCGC ATCCATGCC CCGAACTGCA	2760
GGAGTGGGGA GGCACCGATGG CCGCTTGGT CGAGGCAGGAT CGGGCAGAA ATGGTTGAAC	2820
TCCCGAGAGT GTCCTACACC TAGGGGAGAA GCAGCCAAGG GGTTGTTCC CACCAAGGAC	2880
GACCCGTCTG CGCACAAACG GATGAGCCCA TCAGACAAAG ACATATTCA TCTCTGCTGC	2940
AAACTTGGCA TAGCTCTGCT TTGCCTGGGG CTATTGGGG AAGTTGCGGT TCGTGCTCGC	3000
AGGGCTCTCA CCCTGACTC TTTAATAGC TCTTCTGTGC AAGATTACAA TCTAAACAAT	3060
TCGGAGAACT CGACCTTCCT CCTGAGGCAA GGACCACAGC CAACTTCCTC TTACAAGCCG	3120
CATCGATTTT GTCTTCAGA AATAGAAATA AGAATGCTTG CTAAAAATTA TATTTTACC	3180
AATAAGACCA ATCCAATAGG TAGATTATTA GTTACTATGT TAAGAAATGA ATCATTATCT	3240
TTTAGTACTA TTTTACTCA AATTCAAGAAG TTAGAAATGG GAATAGAAAA TAGAAAGAGA	3300
CGCTCAACCT CAATTGAAGA ACAGGTGCAA GGACTATTGA CCACAGGCCT AGAAGTAAAA	3360
AAGGGAAAAA AGAGTGTGTT TGTCAAAATA GGAGACAGGT GGTGGCAACC AGGGACTTAT	3420
AGGGGACCTT ACATCTACAG ACCAACAGAT GCCCCCTTAC CATATACAGG AAGATATGAC	3480
TTAAATTGGG ATAGGTGGGT TACAGTCAAT GGCTATAAAG TGTTATATAG ATCCCTCCCT	3540
TTTCGTGAAA GACTGCCAG AGCTAGACCT CCTTGGTGTGTA TGTTGTCTCA AGAAGAAAAA	3600
GACGACATGA AACAAACAGGT ACATGATTAT ATTTATCTAG GAACAGGAAT GCACTTTGG	3660
GGAAAGATTG TCCATACCAA GGAGGGGACA GTGGCTGGAC TAATAGAACAA TTATTCTGCA	3720
AAAACTCATG GCATGAGTTA TTATGAATAG CCTTTATTGG CCCAACCTTG CGGTTCCCAG	3780
GGCTTAAGTA AGTTTTGGT TACAAACTGT TCTTAAACAG AGGATGTGAG ACAAGTGGTT	3840
TCCTGACTTG GTTTGGTATC AAAGGTTCTG ATCTGAGCTC TGAGTGTCT ATTTCCTAT	3900
GTTCTTTGG AATTATCCA AATCTTATGT AAATGCTTAT GTAAACCAAG ATATAAAAGA	3960
GTGCTGATTT TTGAGTAAA CTTGCAACAG TCCTAACATT CACCTCTTGT GTGTTGTGT	4020
CTGTTGCCA TCCCGTCTCC GCTCGTCACT TATCCTTCAC TTTCCAGAGG GTCCCCCGC	4080
AGACCCCGGC GACCCTCAGG TCGGGCGACT GCGGCAGCTG GCGCCCGAAC AGGGACCCTC	4140
GGATAAGTGA CCCTGTCTC TATTCTACT ATTTGGTGTGTT TGTCTTGTAT TGTCTCTTTC	4200
TTGTCTGGCT ATCATCACAA GAGCGGAACG GACTCACCAC AGGGACCAAG CTTGTCGACA	4260
TCTAGGGCGG CCAATTCCGC CCCTCTCCCT CCCCCCCCCC TAACGTTACT GGCGAAGCC	4320

GCTTGGAAATA AGGCCGGTGT GCGTTGTCT ATATGTGATT TTCCACCATA TTGCCGTCTT	4380
TTGGCAATGT GAGGGCCCGG AAACCTGGCC CTGTCTTCTT GACGAGCATT CCTAGGGTC	4440
TTTCCCCTCT CGCCAAAGGA ATGCAAGGTC TGTTGAATGT CGTGAAGGAA GCAGTTCCCTC	4500
TGGAAGCTTC TTGAAGACAA ACAACGTCTG TAGCGACCCCT TTGCAGGCAG CGGAACCCCC	4560
CACCTGGCGA CAGGTGCCCTC TGCGGCCAAA AGCCACGTGT ATAAGATACA CCTGCAAAGG	4620
CGGCACAACC CCAGTGCCAC GTTGTGAGTT GGATAGTTGT GGAAAGAGTC AAATGGCTCT	4680
CCTCAAGCGT ATTCAACAAG GGGCTGAAGG ATGCCAGAA GGTACCCCAT TGTATGGGAT	4740
CTGATCTGGG GCCTCGGTGC ACATGCTTTA CATGTGTTTA GTCGAGGTTA AAAAAACGTC	4800
TAGGCCCCCCC GAACCACGGG GACGTGGTTT TCCTTGAAA AACACGCCAA TAATATGGC	4860
GGCCGCTCGA GCATGCATCT AGAACATCGA TAAAATAAA GATTITATTT AGTCTCCAGA	4920
AAAAGGGGGG AATGAAAGAC CCCACCTGTA GGTTGGCAA GCTAGCTTAA GTAACGCCAT	4980
TTTGCAAGGC ATGGAAAAAT ACATAACTGA GAATAGAGAA GTTCAGATCA AGGTCAGGAA	5040
CAGATGGAAC AGCTGAATAT GGGCCAAACA GGATATCTGT GGTAAGCAGT TCCTGCCCG	5100
GCTCAGGGCC AAGAACAGAT GGAACAGCTG AATATGGGC AACACAGGATA TCTGTGGTAA	5160
GCAGTTCCCTG CCCCCGCTCA GGGCCAAGAA CAGATGGTCC CCAGATGCCG TCCAGCCCTC	5220
AGCAGTTCT AGAGAACCAT CAGATGTTTC CAGGGTGCCC CAAGGACCTG AAATGACCT	5280
GTGCCTTATT TGAAACTAACCA AATCAGTTCG CTTCTCGCTT CTGTTCGCGC GCTTCTGCTC	5340
CCCGAGCTCA ATAAAAGAGC CCACAACCCC TCACTCGGGG CGCCAGTCCT CCGATTGACT	5400
GAGTCGCCCG GGTACCCGTG TATCCAATAA ACCCTCTTGC AGTTGCATCC GACTTGTGGT	5460
CTCGCTGTT CTTGGGAGGG TCTCCTCTGA GTGATTGACT ACCCGTCAGC GGGGGCTTT	5520
CATTTGGGGG CTCGTCCGGG ATCGGGAGAC CCCTGCCAG GGACCACCGA CCCACCAACCG	5580
GGAGGTAAGC TGGCTGCCTC GCGCGTTTCG GTGATGACGG TGAAAACCTC TGACACATGC	5640
AGCTCCCGGA GACGGTCACA GCTTGTCTGT AAGCGGATGC CGGGAGCAGA CAAGCCGTC	5700
AGGGCGCGTC AGCGGGTGT GGCGGGTGTG CGGGCGCAGC CATGACCCAG TCACGTAGCG	5760
ATAGCGGAGT GTATACTGGC TTAACTATGC GGCATCAGAG CAGATTGTAC TGAGAGTGCA	5820
CCATATGCGG TGTGAAATAC CGCACAGATG CGTAAGGAGA AAATACCGCA TCAGGCGCTC	5880
TTCCGCTTCC TCGCTCACTG ACTCGCTGCG CTCGGTCGTT CGGCTGCCGC GAGCGGTATC	5940
AGCTCACTCA AAGGCGGTAA TACGGTTATC CACAGAATCA GGGGATAACG CAGGAAAGAA	6000

CATGTGAGCA AAAGGCCAGC AAAAGGCCAG GAACCGTAAA AAGGCCGCGT TGCTGGCGTT	6060
TTTCCATAGG CTCCGCCCCC CTGACGAGCA TCACAAAAAT CGACGCTCAA GTCAGAGGTG	6120
GCGAAACCCG ACAGGACTAT AAAGATACCA GGCGTTTCCC CCTGGAAGCT CCCTCGTGCG	6180
CTCTCCTGTT CCGACCCCTGC CGCTTACCGG ATACCTGTCC GCCTTCTCC CTTGGGAAG	6240
CGTGGCGCTT TCTCATAGCT CACGCTGTAG GTATCTCAGT TCGGTGTAAG TCGTTCGCTC	6300
CAAGCTGGGC TGTGTGCACG AACCCCCCGT TCAGCCGAC CGCTGCGCCT TATCCGGTAA	6360
CTATCGTCTT GAGTCCAACC CGGTAAGACA CGACTTATCG CCACTGGCAG CAGCCACTGG	6420
TAACAGGATT AGCAGAGCGA GGTATGTAAG CGGTGCTACA GAGTTCTTGA AGTGGTGGCC	6480
TAACACTACGGC TACACTAGAA GGACAGTATT TGGTATCTGC GCTCTGCTGA AGCCAGTTAC	6540
CTTCGGAAAA AGAGTTGGTA GCTCTTGATC CGGAAACAA ACCACCGCTG GTAGCGGTGG	6600
TTTTTTGTT TGCAAGCAGC AGATTACGCG CAGAAAAAAA GGATCTCAAG AAGATCCTTT	6660
GATCTTTCT ACGGGGTCTG ACGCTCAGTG GAACGAAAAC TCACGTTAAG GGATTTGGT	6720
CATGAGATTA TCAAAAAGGA TCTTCACCTA GATCCTTTA AATTAAAAAT GAAGTTTTAA	6780
ATCAATCTAA AGTATATATG AGTAAACTTG GTCTGACAGT TACCAATGCT TAATCAGTGA	6840
GGCACCTATC TCAGCGATCT GTCTATTCG TTCACTCCATA GTTGCCTGAC TCCCCGTCGT	6900
GTAGATAACT ACGATAACGGG AGGGCTTACC ATCTGGCCCC AGTGCCTGCAA TGATACCGCG	6960
AGACCCACGC TCACCGGCTC CAGATTATC AGCAATAAAC CAGCCAGCCG GAAGGGCCGA	7020
GCGCAGAAAGT GGTCCCTGCAA CTTTATCCGC CTCCATCCAG TCTATTAATT GTTGCCGGGA	7080
AGCTAGAGTA AGTAGTCGC CAGTTAATAG TTTGCGCAAC GTTGTGCCA TTGCTGCAGG	7140
CATCGTGGTG TCACGCTCGT CGTTGGTAT GGCTTCATTC AGCTCCGGTT CCCAACGATC	7200
AAGGCGAGTT ACATGATCCC CCATGTTGTG CAAAAAAGCG GTTAGCTCCT TCGGTCTCC	7260
GATCGTTGTC AGAAGTAAGT TGGCCGCAGT GTTATCACTC ATGGTTATGG CAGCACTGCA	7320
TAATTCTCTT ACTGTCATGC CATCCGTAAG ATGCTTTCT GTGACTGGTG AGTACTCAAC	7380
CAAGTCATTC TGAGAATAGT GTATGCGCG ACCGAGTTGC TCTTGCCCG CGTCAACACG	7440
GGATAATACC GCGCCACATA GCAGAACTTT AAAAGTGCTC ATCATTGGAA AACGTTCTTC	7500
GGGGCGAAAA CTCTCAAGGA TCTTAACCGCT GTTGAGATCC AGTCGATGT AACCCACTCG	7560
TGCACCCAAAC TGATCTTCAG CATCTTTAC TTTCACCAGC GTTTCTGGGT GAGCAAAAC	7620
AGGAAGGCAA AATGCCGCAA AAAAGGGAAT AAGGGCGACA CGGAAATGTT GAATACTCAT	7680

45

ACTCTTCCTT TTTCAATATT ATTGAAGCAT TTATCAGGGT TATTGTCTCA TGAGCGGATA	7740
CATATTTGAA TGTATTTAGA AAAATAAACCA AATAGGGGTT CCGCGCACAT TTCCCCGAAA	7800
AGTGCCACCT GACGTCTAAG AAACCATTAT TATCATGACA TTAACCTATA AAAATAGGCG	7860
TATCACGAGG CCCTTCGTC TTCAA	7885

**THE EMBODIMENTS OF THE INVENTION IN WHICH AN EXCLUSIVE PROPERTY OR PRIVILEGE IS CLAIMED ARE DEFINED AS FOLLOWS:**

- 1) A method of producing a trans-somatic mammal, wherein said method provides the incorporation of a DNA sequence into the secretory cells of the mammary gland to alter the composition of the milk, wherein said method comprising the steps of:
  - a) providing a vector containing a DNA sequence encoding a valuable compound;
  - b) packaging said vector into a cell line;
  - c) preparing a solution comprising the packaged vector and cell line producing said packaged vector; and
  - d) delivering said solution into the mammary gland to allow the incorporation of the DNA into the secretory cells of the mammary gland.
- 2) The method of Claim 1, wherein the method further comprises the step of flushing the mammary gland with an osmotically-balanced solution prior to delivering the said solution into the mammary gland.
- 3) The method of Claim 2, wherein the method further comprises the step of externally massaging the mammary gland several times a day after delivering the said solution into the mammary gland.
- 4) The method of Claim 2, wherein the method further comprises the step of growing the cells producing said packaged vector on a solid support means, and the solution comprises the cells on said support means and the packaged vector.
- 5) The method of Claim 4, wherein the method further comprises the step of externally massaging the mammary gland several times a day after delivering the said solution into the mammary gland.
- 6) The method of Claim 2 or 4, wherein a substance, with a density higher than the density

of the solution, is delivered into the mammary gland after the delivery of the solution, wherein said substance displaces the solution upwards in the mammary gland.

- 7) The method of Claim 1, wherein the valuable compound is a pharmaceutical.
- 8) The method of Claim 7, wherein the pharmaceutical is a compound selected from the group consisting of: a tissue plasminogen activator, an antibody, an antibiotic, a blood clotting factor, galactosyltransferase, a growth factor, an oncoprotein, a hormone, a milk protein, a hormone receptor, a tumor suppressor protein, a vaccine and an erythropoietin.
- 9) The method of Claim 8, wherein the pharmaceutical is a tissue plasminogen activator.
- 10) The method of Claim 1, wherein the vector is transiently transfected into PA317 cells; the resulting particles are harvested and trans-infected into PG13 cells.
- 11) The method of Claim 1, wherein the vector is selected from the group pL(X)SH, pL(X)SN, pLNS(X), pLHS(X), pLNC(X), pLHC(X), pLNA(X) and pLHA(X); wherein "A" is the beta actin promoter, "L" is the moloney murine virus long terminal repeat (LTR), "S" is the SV40 promoter, "C" is the cytomegalovirus promoter and "X" is a DNA sequence encoding a valuable compound.
- 12) The method of Claim 1, wherein the vector is selected from the group consisting of pL(X)iN and pL(X)iH; wherein "I" is an internal ribosomal entry site (IRES) and "X" is a DNA sequence encoding a valuable compound.
- 13) The method of Claim 1, wherein the vector is pLNM(X); wherein "M" is a mouse mammary tumor virus promoter and "X" is a DNA sequence encoding a valuable compound.
- 14) The method of Claim 1, wherein the vector is pLNMi<sub>2</sub>(X); wherein "M" is a mouse mammary tumor virus promoter, "i<sub>2</sub>" is a modified wild type internal ribosomal entry site and "X" is a DNA sequence encoding a valuable compound.

- 15) The method of Claims 11 to 14, wherein "X" is a DNA sequence encoding a tissue plasminogen activator.
- 16) The method of Claim 1, wherein the solution is an aqueous solution.
- 17) The method of Claim 2, wherein the osmotically-balanced solution is a saline solution.
- 18) The method of Claim 4, wherein the solid support means is a matrix selected from the group consisting of Cytodex beads or Cultisphere.
- 19) The method of Claim 6, wherein said substance is a silicone substance.
- 20) The method of Claim 1, wherein the vector is transfected into a packaging cell line producing a non-retroviral derived particle.
- 21) The method of Claim 1, wherein the vector is transfected into a packaging cell line producing a retroviral derived particle.
- 22) The method of Claim 1, wherein the trans-infecting particle is produced *in vitro*.
- 23) The method of Claim 1, wherein the trans-infecting particle is produced *in vivo*.

1/2

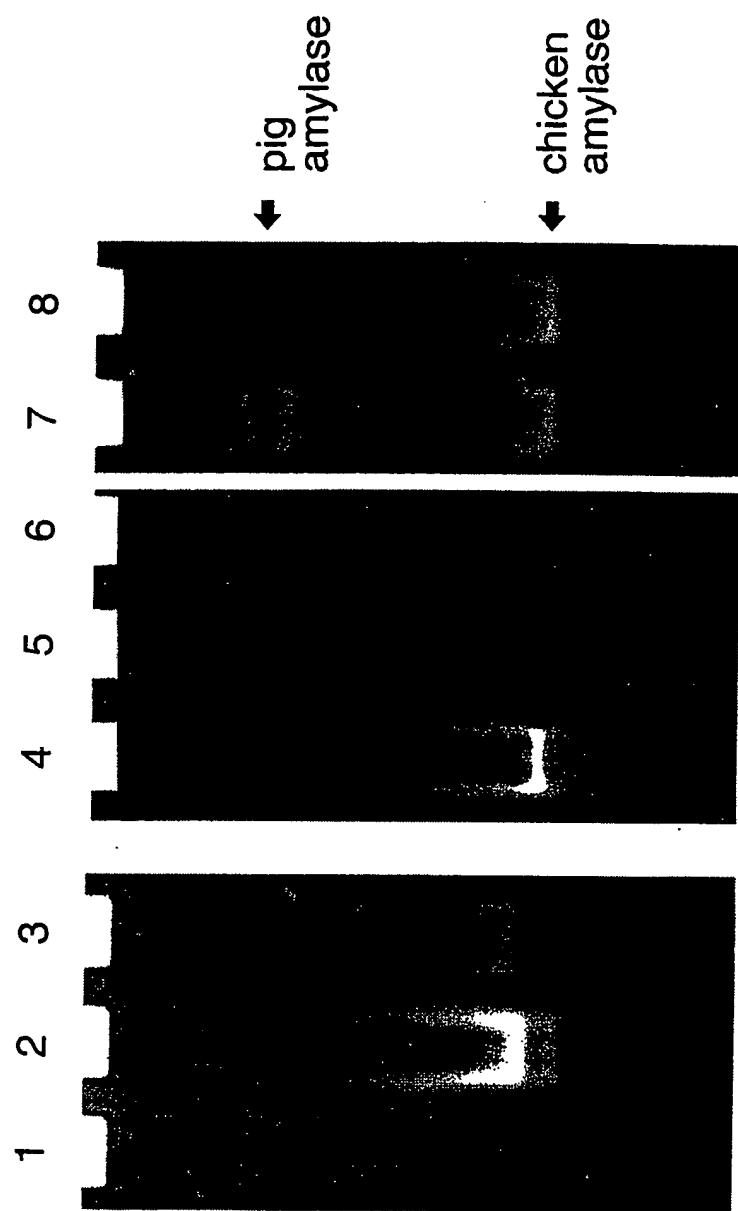
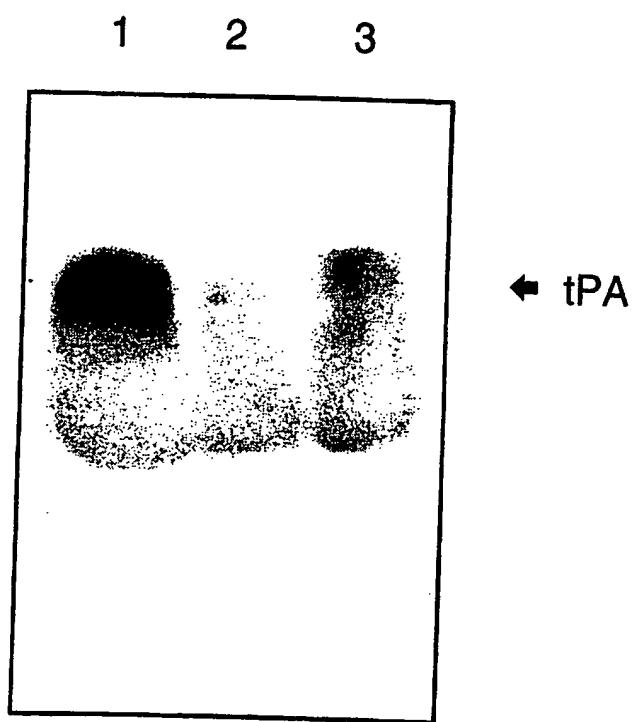


FIG. 1

2/2



**FIG. 2**

# INTERNATIONAL SEARCH REPORT

International Application No  
PCT/CA 98/00607

**A. CLASSIFICATION OF SUBJECT MATTER**  
 IPC 6 C12N15/00 A01K67/027 C12N15/86 C12N9/72

According to International Patent Classification (IPC) or to both national classification and IPC

**B. FIELDS SEARCHED**

Minimum documentation searched (classification system followed by classification symbols)  
 IPC 6 A01K C12N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	ARCHER, J.S. ET AL.: "Human growth hormone (hGH) secretion in milk of goats after direct transfer of the hGH gene into the mammary gland by using replication-defective retrovirus vectors" PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF USA., vol. 91, no. 15, 19 July 1994, pages 6840-6844, XP002088503 WASHINGTON US cited in the application see page 6844, column 1, line 12 - line 14	1,7-11
Y		1-11 -/-

Further documents are listed in the continuation of box C.

Patent family members are listed in annex.

\* Special categories of cited documents :

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier document but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search

Date of mailing of the international search report

18 December 1998

13/01/1999

Name and mailing address of the ISA

Authorized officer

European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl.  
Fax: (+31-70) 340-3016

Chambonnet, F

**INTERNATIONAL SEARCH REPORT**

International Application No

PCT/CA 98/00607

**C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT**

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	YANG, N.S. ET AL.: "Gene transfer into mammalian somatic cells in vivo" CRITICAL REVIEWS IN BIOTECHNOLOGY, vol. 12, no. 4, 1992, pages 335-356, XP002088504 see page 338, column 2 see page 341, paragraph 3. ---	1-11
A	WO 96 22379 A (NEXIA BIOTECHNOLOGIES INC) 25 July 1996 see page 24, line 8 - line 11; claims ---	5,7-9,16
A	US 5 215 904 A (GOULD & WANG) 1 June 1993 cited in the application see the whole document ---	1
Y	CULVER, K.W. ET AL.: "In vivo gene transfer with retroviral vector-producer cells for treatment of experimental brain tumors" SCIENCE., vol. 256, 12 June 1992, pages 1550-1552, XP002088505 LANCASTER, PA US see the whole document ---	1-11
Y	WO 95 15167 A (GENETIC THERAPY INC) 8 June 1995 see the whole document -----	1-11

# INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/CA 98/00607

Patent document cited in search report		Publication date	Patent family member(s)		Publication date
WO 9622379	A	25-07-1996	US	5780009 A	14-07-1998
			AU	4428496 A	07-08-1996
			CA	2210897 A	25-07-1996
			EP	0805869 A	12-11-1997
US 5215904	A	01-06-1993	NONE		
WO 9515167	A	08-06-1995	CA	2177998 A	08-06-1995
			EP	0735887 A	09-10-1996
			JP	9507839 T	12-08-1997

**This Page is Inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record.**

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- BLACK BORDERS**
- IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- FADED TEXT OR DRAWING**
- BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- SKEWED/SLANTED IMAGES**
- COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- GRAY SCALE DOCUMENTS**
- LINES OR MARKS ON ORIGINAL DOCUMENT**
- REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- OTHER:** \_\_\_\_\_

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.**